

# Detection of Low and High Rate DDoS Attack using Metrics with SVM in FireCol Distributed Network

P. Sindhu Priyanka  
PG Scholar

Department of Computer Science and Engineering  
Dr. Mahalingam College of Engineering and  
Technology  
Pollachi – 642003

A. Gowrishankar

M.Tech, Assistant Professor (SG)  
Department of Computer Science and Engineering  
Dr. Mahalingam College of Engineering and  
Technology  
Pollachi – 642003

## ABSTRACT

A federated network mainly operates with same Internet Service Provider (ISP) and virtual entities integrated with it. Foremost frustration in unified network is attack affair due to intruder intervention. Although attacks are classified according to the attack rate dynamics, they are different in many other aspects such as implementations, intention, and countermeasures. Distributed Denial of Service (DDoS) and Low-rate DDoS attacks are vigorous threats to almost every ISP. In a merged network environment, routers work intimately to elevate early warning of DDoS attacks to evade terrible defacement. In existing FireCol a concerted protection, is used to detect flooding attack with metric computations. It delivers better detection for flooding but in case of low-rate attack, with minimum parameters it fights to find. In order to rout that, we prompt additional potential metrics such as Information distance metric, the Generalized entropy metric, the Probability metric ,the Hybrid metric (the Total variation metric and the Bhattacharyya metric) with SVM Classifier for better outsourcing performance in exposure of both high & low transmission rate attacks with diminution in false alarms.

The proportions of packets are being transmitted in distributed client server topology. Both similarity and dissimilarity in the distributions of packets are taken to outline the deviation in the behaviour of user profile. Along with that, SVM classifies the attack and normal flows by using train and test files, which attains the accuracy of 73.89%. Hence, the low rate attack detection with metrics computations and classifier achieves better results compared to Firecol with decision table mechanism.

## General Terms

Attack detection, Security in Distributed network, FireCol, Metric Computations.

## Keywords

DDoS, Flooding attack, Low-rate DDoS attack, Metrics, SVM classifier.

## 1. INTRODUCTION

The Globe has become more consistent with the surfacing of the Internet and the new networking technology. There is a large amount of delicate, viable, social, government and military information need to be protected in the network infrastructure worldwide. Day by day data on network are increasing. So, need of security for the network becomes necessary. The main factors of network security are to maintain integrity, authentication and consistency of network resources. Being resources are distributed; to secure that in

Network & Internet, experts and researchers prefers many attack detection schemes and methods.

Some types of attacks in distributed network are Penetration, Eavesdropping, Man-In-The-Middle, Flooding attacks (DoS) and Distributed Denial of Service (DDoS). Finding solutions to these types of attacks will be focused by attack detection system (so called Intrusion Detection System 'IDS').

In distributed computing, a Denial-of-Service attack (DoS attack) or a Distributed Denial-of-Service attack (DDoS attack) is an attempt to make a machine or network resource unavailable to its deliberate users. It generally consists of efforts to, temporarily/indefinitely interrupt connection or suspend services of a host connected to the network.

To solve these harms, FireCol a collaborative system is taken, which works in ISP's (Internet Service Providers) [2]. An individual Intrusion Prevention System (IPS) or Intrusion Detection system (IDS) can barely sense denial of service attacks, except very close to the user. By permitting massive traffic such as some flooding attacks reach 10–100 GB/s, to passage via the Internet and will be detected/blocked at the host IDS/IPS , which severely strain Internet resources. FireCol, an existing collaborative system which finds high rate(flooding) attacks, far from host and much possible close to the attacking source(s) at the ISP level. This concept will be detailed in ref [2].

Here, detection is mainly focused by calculating the values of various metrics such as 1) Frequency of the traffic [2], 2) Entropy [2], 3) Information distance metric [3], 4) the Generalized entropy metric [3], 5) the Probability metric [9], 6) The hybrid metric (6(a)the total variation metric [9] and 6(b) the Bhattacharyya metric [9] (3,4,5 & 6(a,b) are additionally added as enhanced metrics). Along with these, SVM classifier is a new try to finally trace out the attack accurately. The computed values of the users/clients will be given to train & test with the classifier. This will be detailed in the following sections.

This paper proceeds as follows. Section 2 describes the architecture and global operation of FireCol. The different leveraged metrics and components of the system are presented in Section 3. Section 4 presents FireCol attack detection mechanisms. Section 5 explains the result analysis and discussions paper with outlines of future research directions. Finally, Section 6 explains references.

## 2. EXISTING TECHNIQUE - FIRECOL ARCHITECHTURE

The FireCol structure sustains ring layered protection, and it includes metric manager, the selection manager, score manager and detection window details. A set of IPSs will be in the particular ring. Each FireCol IPS instance will be analyzing aggregated traffic in a detection window details. The frequencies and the entropies of each rule, for each user will be computed by metric manage. A rule could be described as a specific traffic instance to be monitored (might be a traffic filter) which is based on port no's or IP addresses [2]. The following structure (Fig 1) explains the working of FireCol system.

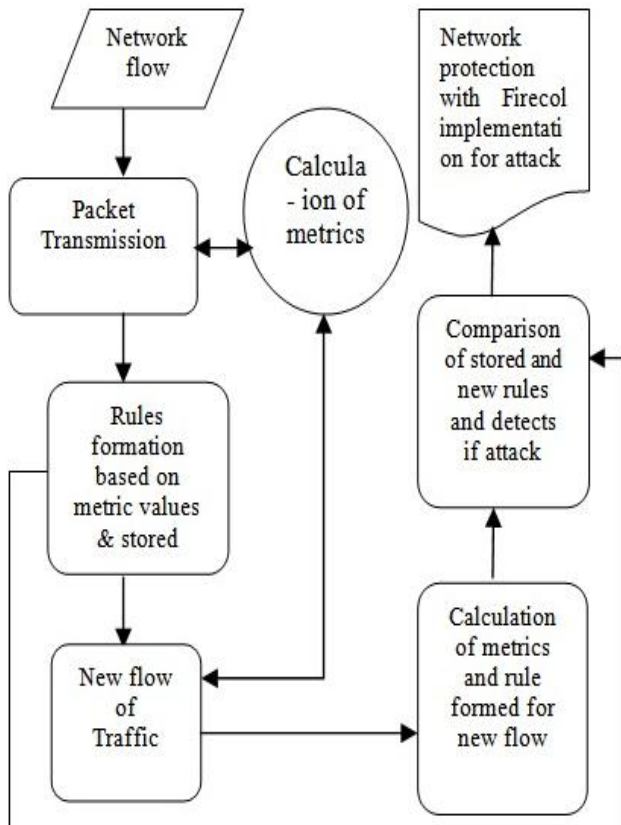


Fig 1: FireCol System Components [Existing]

By following that, selection manager will be deriving the variations in the current distributions from the stored distributions. And the score manager will be assigning the scores for each rule based on the traffic of the particular user, which will be keep on updated with the next upcoming traffics.

The Score detection formula is ,

$$S_i = f_i \times b_j \quad \text{Eqn [1]}$$

Where S-score, f–frequency of particular user and b-threshold value based on decision table.

The architecture diagram for the FireCol collaborative system will be represented as follows in Fig 2,

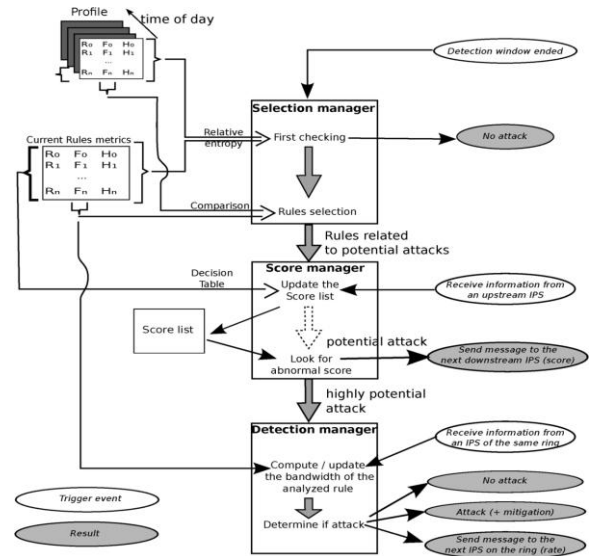


Fig 2: FireCol Architecture [Source: Ref 2]

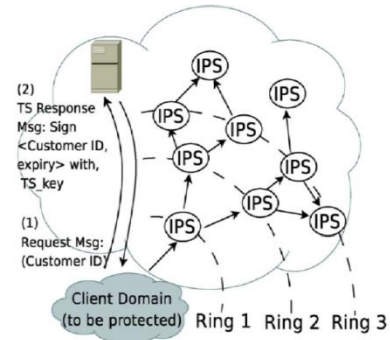


Fig 3: FireCol Subscription [Source: Ref 2]

As shown in the Fig 3, the set of IPSs will be interconnected and the users get registered with it for the protection. But, in this working principle, the rules, scores updates and metrics are mainly concentrated, along with the attention of implementing in the ISP level.

## 3. FIRECOL METRICS WITH NEW ENHANCED METRICS

Mainly, prior to all metrics, rules  $R = \{r_i | i \in [0, n]\}$ , FireCol preserves the values of frequencies and entropies.

### 3.1 Frequency

The Frequency [2]  $f_i$  is the proportion of packets, matching with the rule  $r_i$ , within detection window details, where  $F_i$  is the number of packets matched.

To calculate that, the eqn [1] shown below will be used,

$$f_i = \frac{F_i}{\sum_{j=1}^n F_j} \quad \text{Eqn [2]}$$

The frequency distribution can be defined as,

$$f = \{f_1, \dots, f_n\}. \quad \text{Eqn [3]}$$

### 3.2 Entropy

The entropy [2]  $H$  quantifies the uniformity of distribution of rule frequencies [2]. If all frequencies are identical/same (i.e., uniform distribution), the entropy will be maximal, and the more slanted the frequencies are, the minimal the entropy is.

$$H = -E[\log_n f_i] = \sum_{i=1}^n f_i \log_n (f_i) \text{ Eqn [4]}$$

### 3.3 Information Distance Metric

In Information Distance Metric [3], here consider two discrete complete probability distributions,

$$P = (p_1, p_2, \dots, p_n) \text{ and } Q = (q_1, q_2, \dots, q_n) \text{ with,}$$

$$\sum_{i=1}^n p_i = \sum_{i=1}^n q_i = 1, \quad 1 \geq p_i \geq 0, \quad 1 \geq q_i \geq 0, \quad i = \{1, 2, \dots, n\}. \text{ Eqn [5]}$$

The information divergence is a measure of the divergence between  $P$  and  $Q$  and is shown in the below equation [6],

$$D_\alpha(P||Q) = \frac{1}{\alpha-1} \log_2(\sum_{i=1}^n p_i^\alpha q_i^{1-\alpha}), \quad \alpha \geq 0 \text{ Eqn [6]}$$

Where,  $\alpha$  is a positive arbitrary parameter.

### 3.4 Generalized Entropy Metric

Generalized entropy metric [3], is a measure of the uncertainty allied with a random variable which forms the basis for distance and divergence measurements between victims' various probability densities. The more the variable is random, equal to that; bigger the entropy will be get.

$$H_\alpha(x) = \frac{1}{1-\alpha} \log_2(\sum_{i=1}^n p_i^\alpha) \text{ Eqn [7]}$$

Where,  $P_i$  are the probabilities of  $\{x_1, x_2, \dots, x_n\}$ ,  $P_i \geq 0$ .

### 3.5 Probability Metric

The probability metric [9] is an arithmetic task on the space or distance of distributions of arbitrary elements. It ought to satisfy the following conditions and properties. Consider the probability metric is  $D(x, y), \forall x, y, z \in R$ . It has:

- Identity property:  $D(x, y) = 0$  while  $x = y$ .
- Symmetry property:  $D(x, y) = D(y, z)$ .
- Triangle inequality:  $D(x, y) \leq D(x, z) + D(z, y)$ .

The probability metric contains many classes which include the Hybrid metric (the total variation metric & the Bhattacharyya metric). These both are considered as the most significant probability metrics. The total variation metric and the Bhattacharyya metric mainly compute the difference of two discrete probability distributions and the similarity of two discrete probability distributions respectively [9].

### 3.6 The Total Variation Metric

The total variation [4] is one among the important divergences in mathematical statistics. It can be used measure the largest probable variation between two probability distributions which can assign to the identical event.

Consider here, two complete probability distributions are,

$$P(p_1, p_2, \dots, p_n) \text{ \& } Q(q_1, q_2, \dots, q_n) \text{ with,}$$

$$\sum_{i=1}^n p_i = \sum_{i=1}^n q_i, \quad 1 \geq p_i \geq 0, \quad 1 \geq q_i \geq 0, \quad i = \{1, 2, 3, \dots, n\}.$$

The total variation between them will be as follows:

$$T(P, Q) = \sum_{i=1}^n |p_i - q_i| \quad T(P, Q) = 0, \text{ Eqn [8]}$$

While  $P = Q; T(P, Q) = T(Q, P)$ .

### 3.7 The Bhattacharyya Metric

The Bhattacharyya coefficient [4] is an interesting statistical measure which measures the similarity of two discrete probability distributions and it is also called as similarity coefficient. Here two complete probability distributions  $P$  and  $Q$ , and the conditions will be as in the above metric. The Bhattacharyya coefficient between  $P$  and  $Q$  is denoted by  $\rho(P, Q)$  and its definition as follows:

$$\rho(P, Q) = \sum_{i=1}^n \sqrt{p_i q_i} \text{ Eqn [9]}$$

where,  $\rho$  is Bhattacharyya co-efficient.

The properties of the Bhattacharyya coefficient are explained as follows:

- $0 \leq \rho(P, Q) \leq 1$ ,
- $\rho(P, Q) = 1$  while  $P = Q$
- $\rho(P, Q) = 0$  While  $P$  is orthogonal to  $Q$

The Bhattacharyya coefficient [4] has the symmetric property and is also a probability metric as  $\rho(P, Q) = \rho(Q, P)$ . The value of the Bhattacharyya coefficient indicates the similarity between two probability distributions, unit indicates the strongest similarity, and on the contrary, zero indicates the weakest similarity between the distributions. The following flow diagram Fig 4 explains how the proposed system is going to work with new metrics along with the classifier.

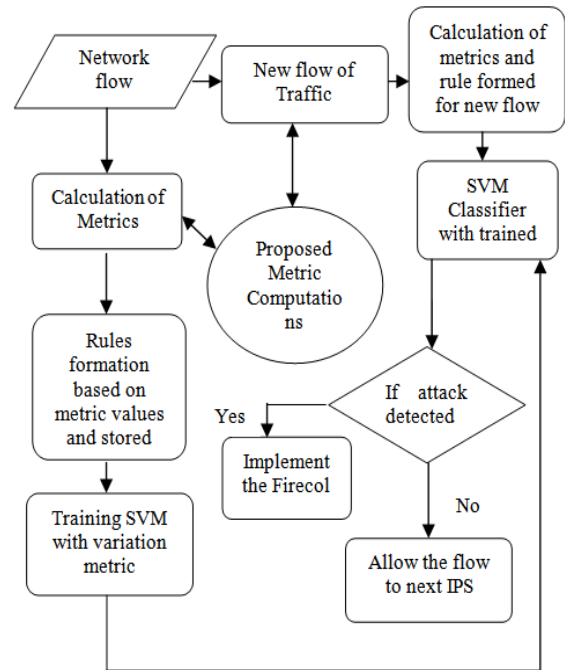


Fig 4: Current system with enhanced proposed metrics

## 4. ATTACK DETECTION METHOD

### 4.1 Training and Testing process using SVM Classifier

Intrusion Detection attempts to spot computer attacks by examining various data records observed in processes on the network. Anomaly discovery has attracted the attention of many researchers to overcome the disadvantage of signature-based IDSs in discovering complex attacks. Although there are some existing mechanisms for Intrusion detection, there is need to improve the performance. Machine Learning techniques are a new approach for Intrusion detection.

The goal of this research is using the SVM machine learning model with different kernels and different kernel parameters for classification of unwanted behavior on the network with scalable performance. First, as shown in Fig 4, the metric computation values will be given to the classifier for training. And then, when a new incoming traffic flow is arrived, it will be analyzed in the testing phase of the classifier to detect the attack patterns.

SVMs consistently outperform other classifiers, in terms of training time and accuracy of detection. So, if the attack pattern is detected via SVM, the particular user will be blocked from the network transmission to shield the network resources. In this examine, by using new potential metric calculations varied from existing scheme with classification technique might increase the detection of the Low-rate DDoS attack. Along with the computation of entropy and frequency, the generalized entropy metric the information distance metrics, the hybrid metric (the total variation metric and the Bhattacharyya metric) could be added for the effective detection. Using these metrics, the similarity as well as the dissimilarity in the network distributions will be calculated.

### 5. RESULT AND DISCUSSION

The computation of various metrics leads to the detection of Low-rate DDoS along with Flooding attack, with the help of traffic including few clients. As discussed above, using decision table & metrics, attackers traced and blocked in existing system.

In existing system, the decision table had been used to detect the attack happened in distributed network. In spite of that, SVM has been focused here to classify Low, High & Normal flows.

Values of Main parameters

$\alpha$	0.4	$\omega$	0.05
$\gamma$	0.8	$\beta$	0.4
$b_1$	1	$b_2$	0.65
$b_3$	0.8	$\tau$	0.5
$\epsilon$	0.01	$\nu$	0.05

Based on these parameters, high rate attack detected with existing three metric values. In proposed methodology, the newly added potential metrics values are recomputed and updating its mean value as per the user flow. Based on that, it would trace out the low rate attack and discriminate it with other attacks as shown in the following snap shots.

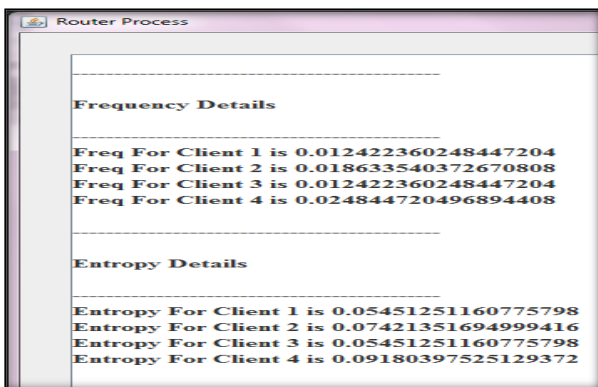


Fig 5: Calculation of Frequency & Entropy

The Fig 5 & 6 represents the values calculated for the few clients in the traffic. The following window shows that the values calculated for relative entropy. The relative entropy metric is necessary because even if two distributions were different, they still can have the same simple entropy.

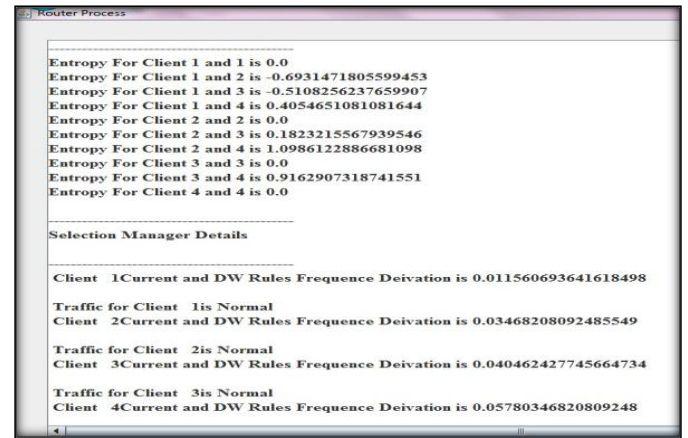


Fig 6: Calculation of Relative Entropy & Score

And the Fig 7 represents the updated instance analyzed aggregated traffic.

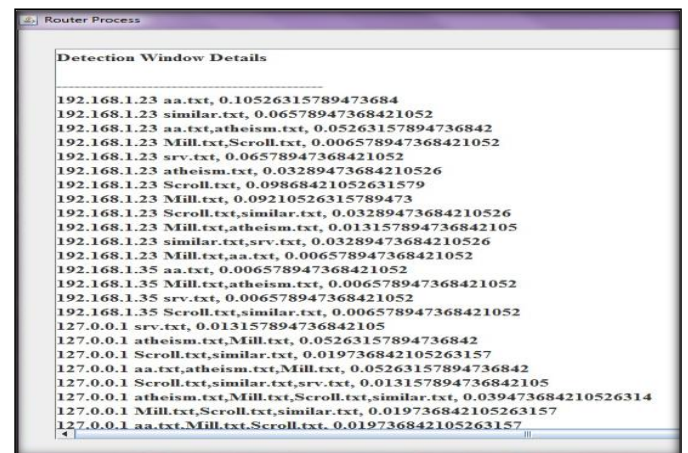


Fig 7: Detection window details

By following that, Fig 8 represents the values calculated for the proposed metrics.

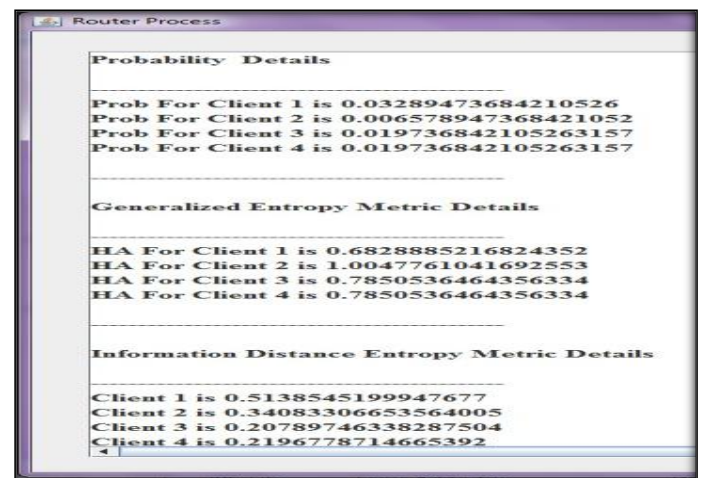


Fig 8(a): Calculation of proposed metrics

Total Variation Metric Details	
Client 1 is	0.5674818216609262
Client 2 is	0.43044967470340606
Client 3 is	0.5036624569460391
Client 4 is	0.47694221201683895
Bhattacharyya Metric Details	
Client 1 is	0.49272730142971727
Client 2 is	0.7773529475071541
Client 3 is	0.6412010254902072
Client 4 is	0.693994272076356
totk	0.49272730142971727
totk	0.7773529475071541
totk	0.6412010254902072
totk	0.693994272076356

Fig 8(b): Calculation of proposed metrics

By having metric computed values, it could be easy to train & test with the SVM classifier to find the deviations and come out results of detecting Low & High rate attacks with effective low false positive rates. Here by as been discussed, it quantified that, as the basis of rule formation the attacker nodes has been found with the metric value deviations and blocked from move onto the next node.

Then, the security in distributed network with the help of metric calculations and SVM classification, against the DDoS attack (such as Flooding & Low-rate attacks) might be attained with improved performance results. This is used to detect and avert the Low-rate DDoS attack, so that sanctuary to network resources, user privacy and privileges might be attained. By using some other classifiers or enhanced rules, numerous types of attacks can be focused to detect in future.

The following snapshots (Fig 9(a), 9(b) & 9(c)) shows that how the packet delivery ratio, throughput and attack (low & high rate) in the current working system respectively.

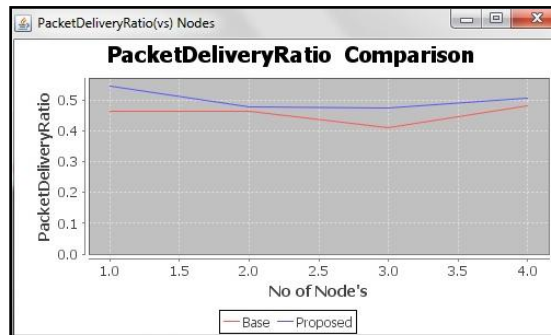


Fig 9(a): PacketDeliveryRatio Comparison

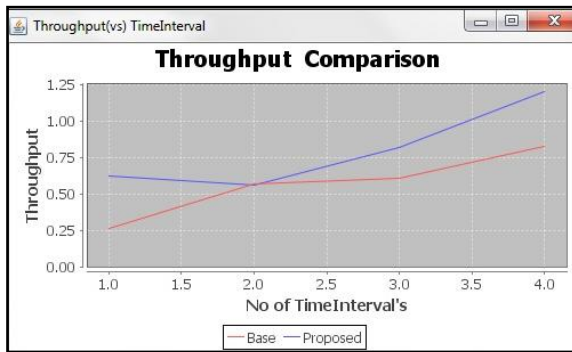


Fig 9(b): Throughput Comparison

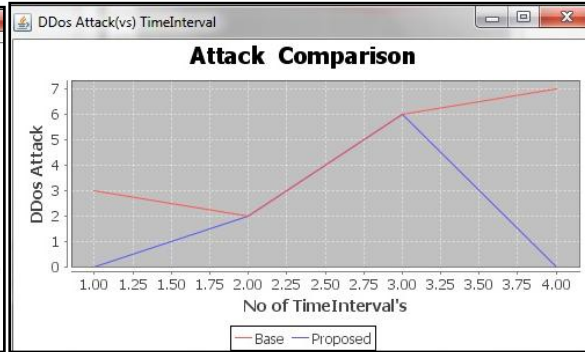


Fig 9(c): Attack Comparison

Clients	InfoMatrix	Gentropy	Frequency	Problity	TotalVariance	Bhattacharyya	Entropy	Attack
1	0.206254136593348...	1.139418697301081	0.003355704697986...	0.003355704697986...	0.5669695096827359	0.3208863177988509	0.019117763377534...	LowRateDDoS
2	0.5138323228513233	1.139418697301081	0.003355704697986...	0.003355704697986...	0.381847960444993...	0.7443718754703137	0.019117763377534...	Normal
3	0.211273523209839...	1.139418697301081	0.003355704697986...	0.003355704697986...	0.4527997527812113	0.6174038733010474	0.019117763377534...	Normal
4	0.219303518953555...	1.139418697301081	0.003355704697986...	0.003355704697986...	0.426052740008240...	0.6681071460160389	0.019117763377534...	Normal

Fig 10(a): Attack Detection

Clients	InfoMatrix	Gentropy	Frequency	Problity	TotalVariance	Bhattacharyya	Entropy	Attack
1	0.352673453454843...	0.7850536464356334	0.019736842105263...	0.019736842105263...	0.366290502793296...	0.2598221044913907	0.077472399319305...	Normal
2	0.3142714181996412	0.8661466680572663	0.013157894736842...	0.013157894736842...	0.320008379888268...	0.378129651402628...	0.056983333424820...	Normal
3	0.476761882890229...	1.0047761041692553	0.006578947368421...	0.006578947368421...	0.2770363128491621	0.4615827895156425	0.033051845531883...	Flooding
4	0.221087272049969...	0.8661466680572663	0.013157894736842...	0.013157894736842...	0.259726256983240...	0.4912097525541795	0.056983333424820...	Normal

Fig 10(b): Attack Detection

Here, along with exiting metrics, the new potential metrics computations are also evaluated with the input as proportion of packets from four clients. The metric values are calculated based on the respective metric formulas and the rules have been formed as per the user traffic (history) in the separate file. Every time the user sending traffic, those values will be calculated and maintained in the training file. And particular current user distribution will be preserved in testing file. As a new methodology has been proposed to detect low rate attack with classifier, the accuracy is mainly focused regarding the detection, which attained as 72% with the [3/4] classification.

## 6. CONCLUSION AND FUTURE SCOPE

The analysis shows that the enhanced detection algorithm is immensely effective in identifying DDoS attack, low rate flow and Normal flow. The existing approaches detect packet flooding DDoS attack based on entropy variation accumulated from network. It conceals sudden increase of legitimate flow as DDoS and denies service to them. So, finer granularity is not obtained. Whereas proposed approach will proactively detect low rate attack from flooding attack and normal flow using various potential metric by observing behavior pattern of packets of legitimate users as well as illegitimate users. The best among the detection depends on d based on recompiling mean value based on behavioral differences between various user flows. This indicates that the detecting accuracy of the current working model is higher than that of the existing model.

To be in spotlight, the attack detection of low-rate, traced out earlier, which represents the technique could not allow the traffic to reach its high level to sprains the network resources from reaching the legitimate users. There is a problem in this detection method that it can be easily evaded by the attacker who tries some other protocols to launch unpredictable DDoS attacks which are currently not defined in this detection system. As an extension of this work, additional metric can be added by analyzing the changing behavior of attackers and dynamically updated database of new users can be built in order to trace out the attacker sources quickly which will increase the earlier detection of attack to a great extent.

So, as a future scope to increase the accuracy in detection other classifiers or different IPS rule structures will be focused.

## 7. REFERENCES

- [1] J. Francois, A. El Atawy, E. Al Shaer, and R. Boutaba (2007) 'A collaborative approach for proactive detection of distributed denial of service attacks,' IEEE, Toulouse, France, Vol. 11, pp. 2-16.
- [2] Jerome François, Issam Aib, and Raouf Boutaba (2012) 'FireCol: A Collaborative Protection Network for the Detection of Flooding DDoS Attacks', IEEE/ACM Transactions on Networking, VOL. 20, NO. 6, pp. 1828 – 1841
- [3] Yang Xiang, Member, IEEE, Ke Li, and Wanlei Zhou (2011) 'Low-Rate DDoS Attacks Detection and Traceback by Using New Information Metrics', IEEE Transactions on Information Forensics And Security, vol.6.
- [4] Shui Yu and Wanlei Zhou (2008) 'Entropy-Based Collaborative Detection of DDOS Attacks on Community Networks', Sixth Annual IEEE International Conference on Pervasive Computing and Communications, pp.568-569.
- [5] Tao Peng, Christopher Leckie, and Kotagiri Ramamohanarao (2007) 'Survey of Network-Based Defense Mechanisms Countering the DoS and DDoS Problems', ACM Computing Surveys, Vol. 39, No. 1, Article 3, pp. 20-2
- [6] Evan Cooke, Farnam Jahanian and Danny McPherson (2005) 'The Zombie Roundup: Understanding, Detecting, and Disrupting Botnets', Proc. SRUTI, USENIX Association Berkeley, CA, USA, pp. 39-44.
- [7] Anja Feldmann, Olaf Maennel, Z. Morley Mao, Arthur Berger and Bruce Maggs (2004) 'Locating Internet routing instabilities', Computer Communication Review, Portland, Oregon, USA, Vol. 34, No. 4, pp. 205-218.
- [8] V. Paxson (1997) 'End-to-end routing behaviour in the Internet,' IEEE/ACM SIGCOMM Computer Communication Review, Vol. 5, No. 5, pp. 601-615.
- [9] Kuai Xu, Zhi-Li Zhang, and Supratik Bhattacharyya (2008) 'Internet traffic behaviour profiling for network security monitoring', IEEE/ACM transactions on Networking, Vol. 16, No. 6, pp. 1241-1252.
- [10] Ke Li, Wanlei Zhou, Ping Li, Jing Hai and Jianwen Liu (2009) 'Distinguishing DDoS Attacks from Flash Crowds Using Probability Metrics', Third International Conference on Network and System Security, pp. 9 – 17.
- [11] Mina Guirguis, Azer Bestavros and Ibrahim Matta (2006) 'On the Impact of Low-Rate Attacks', IEEE International Conference on Communications Vol. 5, pp. 2316 – 2321.
- [12] Xiao-Ming Liu, Gong Cheng, Qi Li, and Miao Zhang (2012) 'A comparative study on flood DoS and low-rate DoS attacks' The Journal of China Universities of Posts

and Telecommunications, Vol. 19, Supplement 1, pp. 116-121.

- [13] D. Muruganandam, Dr.J.Martin Leo Manickam, M.A. Vinoth Kumar (2013) 'Detection and Prevention of Low and High Rate Flooding DDoS Attacks' International journal of advanced scientific and technical research, Issue 3 Vol. 3, pp.187-194.
- [14] Wenke Lee and Dong Xiang (2001) 'Information-Theoretic Measures for Anomaly Detection' IEEE Symposium on Security and Privacy, pp.133-143.
- [15] Wei Wang, Xiaohong Guan, Xiangliang Zhang and Liwei Yang (2006) 'Profiling program behavior for anomaly intrusion detection based on the transition and frequency property of computer audit data' Forty third IEEE Conference on Decision and Control, Vol.1, pp.99-10.
- [16] Barron, A.R., Györfi, L. and Van Ver Meulen, E.C. (1992), 'Distribution Estimation Consistent in Total Variation and in Two Types of Information Divergence' IEEE Transactions on Information Theory, Vol. 38 , pp.1437-1454.