Anomaly based DDoS Attack Detection

Chaitanya Buragohain
Tezpur University
Department of Computer Science & Engineering

Manash Jyoti Kalita
Tezpur University
Department of Computer Science & Engineering

Santosh Singh
Tezpur University
Department of Computer Science & Engineering

Dhruba K. Bhattacharyya
Tezpur University
Department of Computer Science & Engineering

ABSTRACT

Distributed denial-of-service (DDoS) attack poses a serious threat to network security. Several methods have been introduced to reduce the damage. However, most of the methods have been found unable to detect the attack in real-time with high detection accuracy. This paper presents a simple yet effective method to detect DDoS attack for all possible attack scenarios given by Mirkoviac [1] viz constant rate, pulsing rate, increasing rate and sub-group. The proposed method is validated using well known CAIDA dataset.

General Terms

Pattern Recognition, Security.

Keywords

Denial of Service (DOS) Attack, Distributed Denial of Service (DDoS) Attack, Information Gain (IG), Attack Rate, Protocol, Feature Selector (FS)

1. INTRODUCTION

The reduction in processing and best-effort in sending of any packet, malicious or not, was the major concern when the Internet was designed. However, the Internet architecture develops an unregulated network path, which is exploited by most cyber attackers with malicious intention. A denial of service (DDoS) attack is intended to make the resources unavailable to its legitimate users. With the recent evolution of sophisticated attack detection approaches, the traditional single source attacks are now countered easily by most defense mechanisms and the source of these attacks can be easily rebuffed or shut down with improved tracking capabilities. However, with the astounding growth of the Internet during the last decade, an increasingly large number of vulnerable systems are now available to attackers. Attackers can now employ a large number of these vulnerable hosts to launch an attack instead of using a single host, an approach which is not very effective and detected easily. A distributed DoS (DDoS) attack is a large-scale, coordinated attack on the availability of services of a victim system or network resources, launched indirectly through many compromised computers on the Internet called zombies.

The first well-documented DDoS attack appears to have occurred in August 1999, when a DDoS tool called Trinoo was deployed in at least 227 systems, to flood a single University of Minnesota computer, which was knocked down for more than 2 days. The first large-scale DDoS attack took placed on February 2000. On February 7, Yahoo! was the victim of a DDoS attack during which its Internet portal was inaccessible for 3 hours. On February 8, Amazon, Buy.com, cable news network and eBay were all hit by DDoS attacks that caused them to either stop functioning completely or slowed them down significantly.

Along with the evolution of new DDoS attack tools viz Hirak et al [27], many DDoS defense mechanisms have also been proposed. These approaches are of three types depending on their locality of deployment: source-end approach, victim-end approach and in-network approach. Detecting any DDoS attack at the victim end is easy, but often not useful after legitimate clients have been denied access. Source-end detection is a very challenging task. The existing detection approaches can be categorized into statistical, soft computing, clustering, knowledge-based and hybrid. These approaches can also be classified as super-vised or unsupervised [2]. Clustering is a data mining technique, which is also known as unsupervised classification. It does not require a training dataset and the strength of clustering lies within the algorithm itself. Hence, it is very popular. Classifiers such as support vector machine (SVM) and hidden Markov model are also used in many detection approaches. A detailed discussion on these approaches and methods can be found in [28].

Based on the limited survey, following are the observations: The detection approaches for DDoS attack based on their deployability can be of three types: source-end, victim-end and intermediate router defense. Most existing methods belong to victim-end category. However, two of the most essential requirements of this category of methods are :(i) Real-time or near real-time detection and (ii) High detection accuracy.

Existing methods can be categorized based on the approach used for detection such as statistical, knowledge-based, soft-computing, other data-mining and machine-learning methods etc. Most methods have been found failing to detect in real time due to their less cost effective mechanism (use of too many parameters or complex detection logic).

Due to non-availability of standard intrusion datasets, proper validation of the methods has been a major bottleneck. Existing DDoS attack scenarios can be of four types: pulsing rate, constant rate, increasing rate and sub-group.

This paper introduces an effective method for DDoS attack detection using standard approach for all possible attack scenarios in real-time. The method has been established to perform satisfactorily using a benchmark dataset.

The remainder of the paper is organized as follows. Section 2 presents some reviews and related works. Section 3 introduces the proposed architecture and algorithm. Simulation results are discussed in section 4. Section 5 concludes the paper.

2. RELATED WORK

In this section, a selected review on existing literature on DDoS attack detection methods is presented. Recent trends show that soft-computing approaches have been used heavily for DDoS attack detection. The ensembles of classifiers have also performed satisfactorily with high detection rates (DRs). The methods for DDoS attack detection can be classified into four major classes, as reported below.

2.1 Statistical Methods

This category of methods attempt to identify DDoS attack using statistical approaches, either in centralized or in a distributed mode. Most of these methods [1], [3]-[8] adopt victim-end approach and attempt to handle the attack with minimum number of features.

2.2 Soft Computing Method

In this category, the developer attempts to identify DDoS attack based on various soft computing techniques such as Artificial Neural Network, Fuzzy reasoning etc. Such methods [9]-[13] are mostly supervised in nature and perform well for known attacks. However, lack of an appropriate training dataset often becomes a major bottleneck for these methods.

2.3 Knowledge Based Model

A knowledge-based method can be deployed in the victim side source network as well as intermediate network. Similarly such methods [14]-[18] works based on packet as well as flow information.

2.4 Other Data Mining and Machine learning Methods

Several researchers have attempted to handle DDoS attack using other data mining and machine learning approaches [19]-[26], [29].

2.5 Discussion

Based on the theoretical as well as experimental study it has been observed that,

1. Statistical methods are capable of detecting attacks from normal traffic if there exist distinct statistical properties in both classes of traffic.
2. Soft computing methods can handle DDoS attack with high detection accuracy but often found less cost effective. The success of such methods is dependent on user input parameters which are crucial from the system as well as user point of view.
3. Knowledge-based methods are supervised in nature, which perform satisfactorily both from detection accuracy as well as real-time performance point of view provided prior knowledge is available.
4. Data mining and machine learning methods can be both supervised as well as unsupervised in nature are capable of detecting DDoS attacks with high detection accuracy; however, such methods fail to perform in real-time.
5. Most researchers are in favor of using the victim-end approach. But, a common disadvantage of this scheme is the consumption of a huge amount of resources to provide a fast detection response. In the latest DDoS attack scenarios, once the attackers gain access, they can increase the attack intensity instantly, and after acquiring a majority of available resources, they can launch attacks without spoofing IPs and can extend activities to complex database query transactions. Generally, segregation of such transactions from the legitimate queries is a difficult task.

3. PROPOSED WORK

Current DDoS attacking tools are capable of launching attacks in different modes [15] such as increasing rate, constant rate, pulsing rate (attack rate oscillates between maximum to 0) and gradual pulsing (e.g. attack rate achieves a maximum in 20 s and reduces to 0 in 10 s). Here an algorithmic approach is presented which generates an alarm when any sort of abnormality occurs. After the alarm is generated the kind of attack (protocol specific) that has occurred is distinguished. The proposed method works in four major steps viz., information gain based feature ranking, rate analysis, packet analysis and protocol analysis.

3.1 Feature Selection

The proposed detection method uses an ensemble of feature selectors using PCA based feature selector and correlation based feature selector. A weighted majority based voting is used to combine the output of each individual feature selector. The weight of each feature selector (FS) is decided on the individual performance of each FS. In this proposed experimentation, it has been observed that the performance of information gain (IG) based on FS is best.

3.2 Proposed Architecture for Network Traffic Analyzer

The raw data from the network is first sent to the capture and pre-processing unit of the analyzer. Attributes are selected based on various protocols like TCP, UDP, and ICMP etc. The data rate for the entire data is calculated on the basis of time-stamp. After the rate is calculated, various attacks are distinguished on the basis of thresholds obtained through information gain approach for which the packets are classified as anomalous or normal. A conceptual framework of the proposed method is shown in the fig 1. The method can work in two levels i.e. (i) raw traffic level and (ii) preprocessed traffic feature level. In case of raw traffic level detection, the method selects a minimum number of attributes (e.g. timestamp, Source IP, Destination IP, protocol etc.) and attempts to identify all possible DDoS attack scenarios using a faster packet rate analyzer algorithm over a t-second window interval.

3.3 Algorithm

Here a statistical approach towards the network analysis based on deviation from the standard behavior of the network traffic is presented. The Algorithm has been divided into three parts: Rate Analyzer, Packet Analyzer and Protocol Analyzer.

3.3.1 Rate Analyzer

For each t-second interval (consider k=5 sec), the time stamp attribute parameter of each packet is read and the difference between the time stamps is calculated. If the calculated value is equal to the given time interval, the corresponding rate is obtained. If not then the calculation process is continued. The pseudo code of Rate Analyzer is presented in Algorithm1.

Algorithm 1: Rate Analyzer

Input: c: CAIDA dataset, k: time window size, t: time stamp

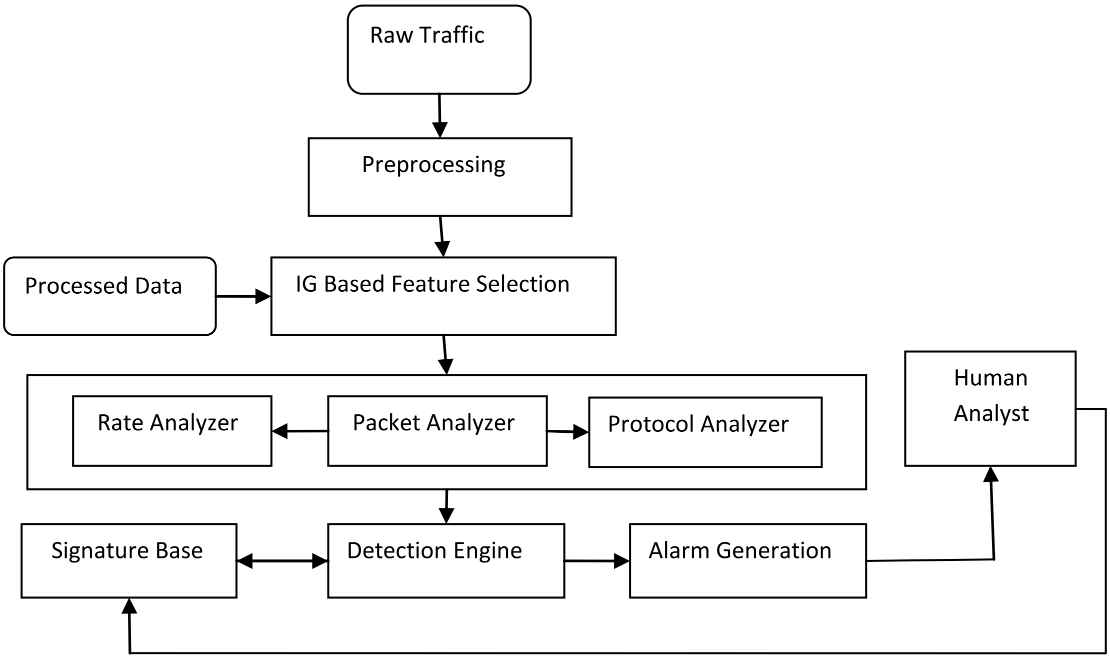

Fig 1: Proposed architecture for network traffic analyzer

Output: time frames and packet rates

```
procedure Rate Analyzer
  while End of File not reached do
    for each incoming packet do
      read t
      find diff of t
      if diff  = k then
        rate= packetcount/k
        break
      else
        packetcount++
        continue
      end if
    end for
end procedure
```

3.3.2 Packet Analyzer

The algorithm uses six user defined parameters to analyze the packets involved in network traffic. If the packet rate (calculated in Rate Analyzer) is greater than a user defined threshold ϴ, there is a chance of occurrence of an attack. Also if the difference obtained is negligible then an attack has occurred and the pattern shown by them is that of constant attack. If the difference is greater than α and less than β, then the attack obtained is that of linearly increasing pattern. If the difference obtained is greater than α and less than ϒ, then the given attack pattern is that of linearly or exponentially increasing pattern. Again if the difference of the time stamps fluctuates for a given period then such pattern is considered as pulsing attack. Else if neither of the above patterns is obtained, then the traffic pattern is considered as normal. The pseudo code of Packet Analyzer is presented in Algorithm 2.

Algorithm 2: Packet Analyzer

Input: time frames, packet rates,

ϴ: threshold for packet rates,

ʅ: threshold for constant attack,

α: minimum threshold for increasing attack,

β: maximum threshold for linear increasing attack,

ϒ: maximum threshold for exponential increasing attack,

∂: threshold for pulsing behavior

Output: various forms of attacks

```
procedure Packet Analyzer
  while End of File not reached do
    if packet rate> ϴ then
      if diff < ʅ  then
        if End of File reached then
          Constant attack scenario
          break
        else
          continue
        end if
      else if diff > α  && diff < β  then
        if End of File reached then
          Linearly increasing attack scenario
          break
```

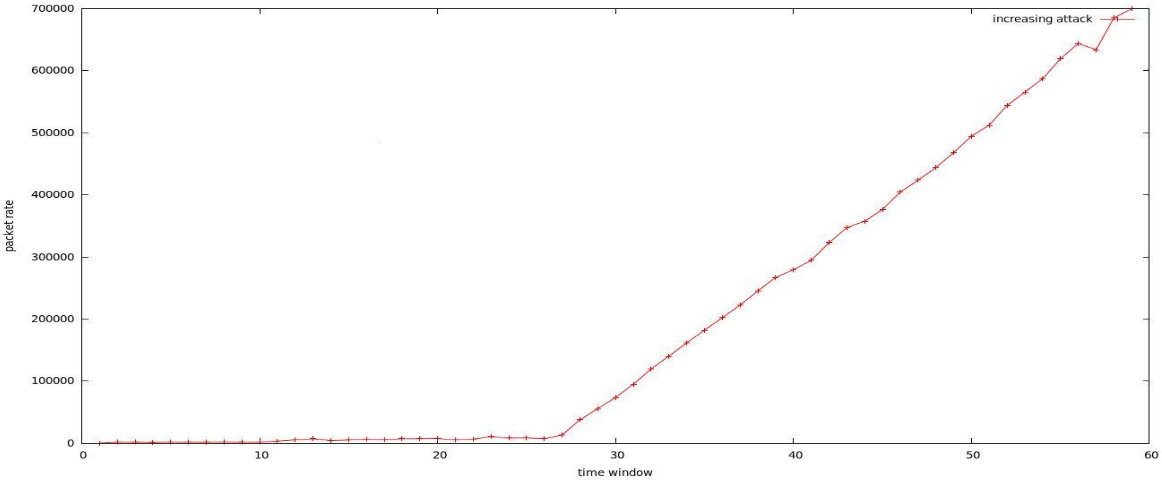

Fig 2: A figure for increasing attack

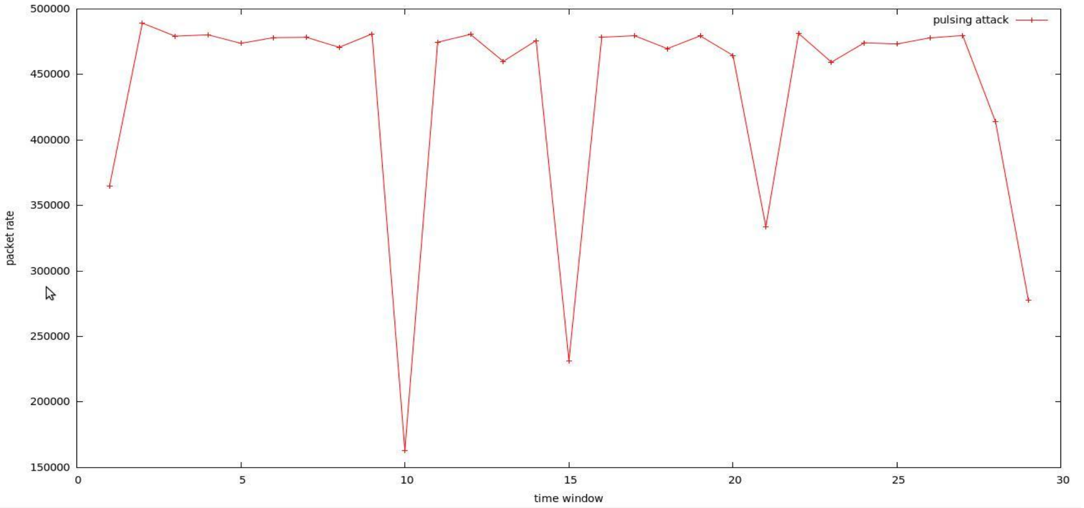

Fig 3: A Figure for Pulsing Attack

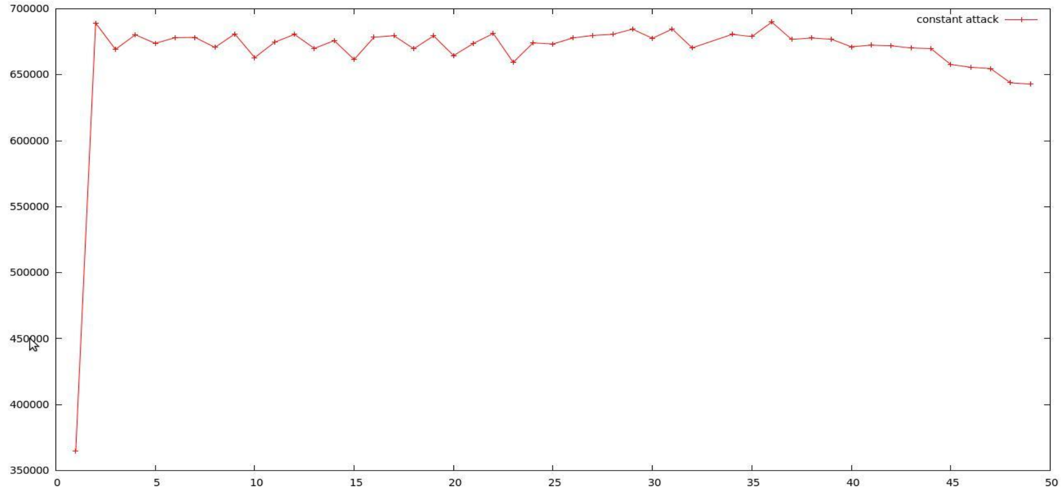

Fig 4: A Figure for Constant Attack

```
else
    continue
end if
else if diff > $\alpha$ && diff < $\Upsilon$ then
    if End of File reached then
        Linearly or Exponentially increasing attack scenario
        break
    else
        continue
    end if
else if diff < $\alpha$ then
    if End of File reached then
        break
```

```
else
continue
end if
else
flag++
continue
end if
if flag> ∂ then
Pulsing attack scenario is generated
else
Normal traffic scenario
end if
end procedure
```

3.3.3 Protocol Analyzer

The algorithm given in this subsection analyses the packets in terms of their protocol. This mainly finds what type of protocol or attack is involved during the anomaly traffic. The pseudo code of Protocol Analyzer is presented in Algorithm 3.

Algorithm 3: Protocol Analyzer

Input: Ω: threshold value, c: Caida dataset, t: time window size

Output: packets of different protocol

```
procedure Protocol Analyzer
if diff = k then
rate= packetcount/k
 break
if number of packets > Ω then
while End of File not reached do
if packet= ARP then
arp++
else if packet=ICMP then
icmp++
else if packet=TCP then
tcp++
else if packet=UDP then
 udp++
else if packet=ACK then
ack++
end if
else
 packetcount++
 Continue
end if
end procedure
```

4. EXPERIMENTAL RESULTS

We have performed our experiment on CAIDA dataset. There are various types of attack scenarios present in the dataset. By using the proposed algorithmic approach attacks like Increasing, Pulsing and Constant are obtained. Three different graphs are plotted by using GNU plot for the above three different attacks. They are depicted in fig 2, fig 3 and fig 4.

4.1 Discussion

In this paper an attack detection method is presented which helps in distinguishing various attack scenarios using a statistical approach, more precisely a method based on information gain. The attacks are first divided into various categories using threshold for each one. These thresholds are based on experimental results. The graphs plotted, using GNU plot, gives an avid picture of various attacks encountered and categorization of the rate of attacks and the protocol available in each packet for attack launching. Based on the protocols it distinguishes how many packets of each protocol are available in the given attack. The proposed method is sensitive to threshold parameters. To obtain the mentioned results an exhaustive experiment is conducted heuristically. Thus using this approach the required output can be obtained.

The execution time for the proposed algorithmic approach is around eighty seconds (80 seconds) which is comparatively less than the total time of the attack scenario on which the experiment was performed. So it can be concluded that this algorithmic approach is a near real-time one.

Table 1. Table of detection accuracy

	False Alarm		**Detection**	
	Normal Data	Attack Data	Normal Data	Attack Data
CAIDA Data	0.05	0.01	0.01	0.93
Simulated Network	0.06	0.06	0.93	0.95

5. CONCLUSION AND FUTURE WORK

This paper presents an overview of DDoS attack, detection schemes and finally proposed a method to detect various attack patterns. Most of the methods for DDoS anomaly detection are either not too effective, not accurate or are complex in nature and firm to implement. But the method proposed in this paper is very simple in concept and easy to implement. The method has been implemented in a simulated environment and the results are found to be satisfactorily accurate. The effectiveness of the method has been established using the well-known CAIDA dataset. Work is going on to extend the present work to identify DDoS attack using a distributed approach.

6. REFERENCES

[1] Mirković, J., Gregory, P. and Peter, R. 2002. Attacking DDoS at the source. In Proceedings of the 10th IEEE International Conference on Network Protocols.

[2] Specht, S. M. and Ruby B. L. 2004. Distributed Denial of Service: Taxonomies of Attacks, Tools, and Countermeasures. In Proceedings of the 17th International Conference on Parallel and Distributed Computing Systems.

[3] Chen, Y., Hwang, K. and Ku W. S. 2006. Distributed change-point detection of DDoS attacks over multiple network domains. In Proceedings of the IEEE International Symposium on Collaborative Technologies and Systems.

[4] Chen, C. L. 2009. A New Detection Method for Distributed Denial-of-Service Attack Traffic based on Statistical Test. Journal of Universal Computer Science.

[5] Akella, A., Bharambe, A., Reiter, M. and Seshan, S. 2003. Detecting DDoS attacks on ISP networks. In Proceedings of the Workshop on Management and Processing of Data Streams, ACM.

[6] Öke, G. and Loukas. G. 2007. A denial of service detector based on maximum likelihood detection and the random neural network. The Computer Journal.

[7] Cheng, J., Yin, J., Wu, C., Zhang, B. and Li, Y. 2009. DDoS attack detection method based on linear prediction model. In Proceedings of the 5th inter-national conference on Emerging intelligent computing technology and applications.

[8] Udhayan, J. and Hamsapriya, T. 2011. Statistical segregation method to minimize the false detections during DDoS attacks. International Journal of Network Security.

[9] Nguyen, H.V. and Choi, Y. 2010. Proactive detection of DDoS attacks utilizing k-NN classifier in an AntiDDoS framework. International Journal of Electrical, Computer, and Systems Engineering.

[10] Shanon, C. E. 1948. A mathematical theory of communication. Bell system technical journal.

[11] Gavrilis, D. and Dermatas, E. 2005. Real-time detection of distributed denial-of-service attacks using RBF networks and statistical features. Computer Networks and ISDN System.

[12] Wu, Y. C., Tseng, H. R., Yang, W. and Jan, R. H. 2011. DDoS detection and traceback with decision tree and grey relational analysis. International Journal of Ad-Hoc and Ubiquitous Computing.

[13] Karimazad, R. and Faraahi, A. 2011. An anomaly based method for DDoS attacks detection using rbf neural networks. In Proceedings of the International Conference on Network and Electronics Engineering.

[14] Jeyanthi, N. and Iyengar, N. C. S. N. 2012. An entropy based Approach to detect and distinguish DDoS aatacks from ash crowds in VoIP networks. International Journal of Network Security.

[15] Thomas, R., Mark, B., Johnson, T. and Croall, J. 2003. NetBouncer: Client-legitimacy-based high performance DDoS filtering. In Proceedings of the 3rd DARPA Information Survivability Conference and Exposition.

[16] Limwiwatkul, L. and Rungsawang, A. 2004. Distributed denial of service detection using TCP/IP header and traffic measurement analysis. In Proceedings of the IEEE International Symposium Communications and Information Technology.

[17] Zhang, G. and Parashar, M. 2006. Cooperative defense against DDoS attacks. Journal of Research and Practice in Information Technology.

[18] Wang, J., Phan, R. C. W., Whitely, J. N. and Parish, D. J. 2010. Augmented attack tree modeling of distributed denial of services and tree based attack detection method. In Proceedings of the 10th IEEE International Conference on Computer and Information Technology.

[19] Hwang, K., Dave, P. and Tanachaiwiwat, S. 2003. NetShield: Protocol anomaly detection with data-mining against DDoS attacks. In Proceedings of the 6th International Symposium on Recent Advances in Intrusion Detection.

[20] Li, L. and Lee, G. 2003. DDoS attack detection and wavelets. In Proceedings of the 12th International Conference on Computer Communications and Networks.

[21] Sekar, V. Duffield, N., SpatsCheck, O., van der Merwe, J. and Zhang, H. 2006. Lads: large-scale automated DDoS detection system. In Proceedings of the annual conference on USENIX Annual Technical Conference.

[22] Erol, G. and Loukas, G. 2007. A self-aware approach to denial of service defence. Journal of Computer Networks: The International Journal of Computer and Telecommunications Networking.

[23] Lee, K., Kim, J., Kwon, K. H., Han, Y. and Kim, S. 2008. DDoS attack detection method using cluster analysis. Journal of Expert Systems with Applications.

[24] Li, M. and Li, M. 2009. A new approach for detecting DDoS attacks based on wavelet analysis. In Proceedings of the 2nd International Congress on Image and Signal Processing.

[25] Dainotti A., Pescapé, A. and Ventre, G. 2009. A cascade architecture for DoS attacks detection based on the wavelet transform. Journal of Computer Security.

[26] Xia, Z., Lu, S. and Li, J. 2010. Enhancing DDoS flood attack detection via intelligent fuzzy logic. Informatica: An International Journal of Computing and Informatics.

[27] Kashyap, H. J. and Bhattacharyya, D. K. 2012. A DDoS attack detection mechanism based on protocol specific traffic features. In Proceedings of the 2nd International Conference on Computational Science, Engineering and Informational Technology.

[28] Bhattacharyya, D. K. and Kalita, J. K. 2013. Network anomaly detection from machine learning perspective. A chapter and hall book, CRC press, Taylor and Francis group.

[29] Rahmani H., Sahli, N. and Kammoun, F. 2009. Joint entropy analysis model for DDoS attack detection. In Proceedings of the 5th International Conference on Information Assurance and Security.