Detecting Distributed Denial of Service Attacks: Methods, Tools and Future Directions

ARTICLE in THE COMPUTER JOURNAL · MARCH 2013
Impact Factor: 0.79 · DOI: 10.1093/comjnl/bxt031

CITATIONS
17

READS
2,537

4 AUTHORS, INCLUDING:

Dhruba Kumar Bhattacharyya
Tezpur University
122 PUBLICATIONS 489 CITATIONS

Jugal Kalita
University of Colorado Colorado Springs
121 PUBLICATIONS 814 CITATIONS

All in-text references underlined in blue are linked to publications on ResearchGate, letting you access and read them immediately.

Available from: Dhruba Kumar Bhattacharyya
Retrieved on: 19 February 2016
Detecting Distributed Denial of Service Attacks: Methods, Tools and Future Directions

Monowar H. Bhuyan¹, H. J. Kashyap¹, D. K. Bhattacharyya¹ and J. K. Kalita²

¹Department of Computer Science & Engineering, Tezpur University, Napaam, Tezpur-784028, Assam, India
²Department of Computer Science, University of Colorado at Colorado Springs, CO 80933-7150, USA
Email: {mhb,dkb}@tezu.ernet.in, jkalita@uccs.edu

Distributed Denial of Service attack is a coordinated attack, generally performed on a massive scale on the availability of services of a target system or network resources. Due to the continuous evolution of new attacks and ever-increasing number of vulnerable hosts on the Internet, many DDoS attack detection or prevention mechanisms have been proposed. In this paper, we present a comprehensive survey of DDoS attacks, detection methods and tools used in wired networks. The paper also highlights open issues, research challenges and possible solutions in this area.

Keywords: DDoS, Denial of Service, Agents, Handler, Network security

1. INTRODUCTION

The minimal processing and best-effort forwarding of any packet, malicious or not, was the prime concern when the Internet was designed. This architecture creates an unregulated network path, which can be exploited by any cyber attacker motivated by revenge, prestige, politics or money. Denial-of-service (DoS) attacks exploit this to target critical Web services [1, 2, 3, 4, 5]. This type of attack is intended to make a computer resource unavailable to its legitimate users.

Denial of service attack programs have been around for many years. Old single source attacks are now countered easily by many 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 server, an approach which is not very effective and detected easily. A distributed denial of service (DDoS) attack [1, 6] 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.

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 two days¹. The first large-scale DDoS attack took place on February 2000¹. On February 7, Yahoo! was the victim of a DDoS attack during which its Internet portal was inaccessible for three hours. On February 8, Amazon, Buy.com, CNN and eBay were all hit by DDoS attacks that caused them to either stop functioning completely or slowed them down significantly¹.

DDoS attack networks follow two types of architectures: the Agent-Handler architecture and the Internet Relay Chat (IRC)-based architecture as discussed by [7]. The Agent-Handler architecture for DDoS attacks is comprised of clients, handlers, and agents (see Figure 6). The attacker communicates with the rest of the DDoS attack system at the client systems. The handlers are often software packages located throughout the Internet that are used by the client to communicate with the agents. Instances of the agent software are placed in the compromised systems that finally carry out the attack. The owners and users of the agent systems are generally unaware of the situation. In the IRC-based DDoS attack architecture, an IRC communication channel is used to connect the client(s) to the agents. IRC

¹http://www.garykessler.net/library/ddos.html
ports can be used for sending commands to the agents. This makes DDoS command packet more untraceable. Moreover, it is easier for an attacker to hide his presence in an IRC channel as such channels tend to have large volumes of traffic. A recent attacking tool by Anonymous based on the IRC protocol is LOIC (Low Orbit Ion Cannon) [8]. It includes three primary methods of attacks for TCP, UDP, HTTP and is found in two versions: binary and web-based. It allows clients to connect remotely via the IRC protocol and to be a part of a system of compromised hosts. The bigger the size of compromised hosts, the more powerful the attack is. Between these two architectures, the agent handler architecture is commonly found in use in the literature.

Along with the evolution of new DDoS attack tools, 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. Detection approaches used include statistical, soft-computing, clustering, knowledge-based and classifiers. These approaches can also be classified as supervised or unsupervised [9].

Statistical techniques fit a statistical model to the given data and then apply a statistical inference test on an unseen instance to determine if it belongs to this model. In knowledge-based methods, predefined rules or patterns of attack are checked against connection events to test their legitimacy. Soft Computing techniques apply problem solving technologies such as fuzzy logic, probabilistic reasoning, neural networks and genetic algorithms. Clustering is a data mining technique, which is also known as unsupervised classification. It does not need to be trained with a training dataset and the strength of clustering lies within the algorithm itself. Hence, it is very popular. Classifiers such as SVM and HMM are also used in many detection approaches. Detailed discussion of these approaches can be found in [9].

In the past few years, several significant surveys have been published [1, 2, 3, 6, 7, 10, 11] to highlight the architectures and methods developed for network defense mechanisms, attack taxonomies, attack launching mechanisms, and their pros and cons. However, our survey differs significantly from them in the following ways.

- We present an attack taxonomy based on [10]. However, in our taxonomy there are seven distinct possibilities in which an intruder can attempt to launch DDoS attacks. Unlike [10], we include a detailed discussion of various DDoS defense mechanisms and their architectures and methods under the broad categories of statistical, knowledge based, soft computing, data mining, and other machine learning methods. We also include a list of practical issues and research challenges, which is not available in [10].
- Like [1], we report and analyze a large number of defense mechanisms, architectures, methods, tools and solutions countering Denial of Service of attacks. However, unlike [1], we include recent defense solutions and tools and discuss latest DDoS attack strategies. Also, unlike [1], we attempt to provide a possible solution to counter the attacks in the context of latest DDoS attack scenarios.
- Unlike [3], our survey is focused on DDoS attack detection methods, tools and research directions. In [3], a major portion is dedicated to DoS research solutions only and that too for a period upto 2009. Also, unlike [3], we present several research issues and open challenges from a more practical point of view, considering the latest DDoS attack scenarios.
- Like [6], we present possible solutions for DDoS attacks in detail and with a practical viewpoint. Also, we broadly categorize the detection methods as statistical, knowledge based, soft computing, data mining and other machine learning. We also include tools related to DDoS attacks.
- Only a brief overview of DDoS taxonomies, tools and countermeasures is given in [7] without any possible solutions for DDoS attacks. In contrast, we present a list of methods, tools, possible solutions and future research directions for DDoS attacks in detail.
- In [2], the authors present several anomaly detection techniques w.r.t. diverse domains but our work is mainly focused on DDoS attack detection architectures, methods, and tools.

Our survey begins in Section 2 with the introduction of DDoS attacks and generic architectures of DDoS defense mechanisms classified with their locality of deployment. Section 3 discusses various methods for DDoS attack detection. Different strategies to evaluate performance of DDoS attack detection methods are described in Section 4. The challenges faced by DDoS defenders are reported in Section 5 followed by concluding remarks in Section 7.

2. DDoS ATTACKS AND THEIR ARCHITECTURES

As stated in [12], a DDoS attack can be defined as an attack which uses a large number of computers to launch a coordinated DoS attack against a single machine or multiple victim machines. Using client/server technology, the perpetrator is able to multiply the effectiveness of the DoS attack significantly by harnessing the resources of multiple unwitting accomplice computers, which serve as attack platforms. Approximate attack statistics for DDoS [11] for the year 2011 are shown in Figure 1. A DDoS attacker is considered more intelligent than a DoS attacker. It is
DDoS attacks mainly take advantage of the architecture of the Internet and this is what makes them powerful. While designing the Internet, the prime concern was to provide for functionality, not security. As a result, many security issues have been raised, which are exploited by attackers. Some of the issues are given below.

- Internet security is highly interdependent. No matter how secure a victim’s system may be, whether or not this system will be a DDoS victim depends on the rest of the global Internet [13, 14].
- Internet resources are limited. Every Internet host has limited resources that sooner or later can be exhausted by a sufficiently large number of users.
- Many against a few: If the resources of the attackers are greater than the resources of the victims, the success of the attack is almost definite.
- Intelligence and resources are not collocated. Most intelligence needed for service guarantees is located at end hosts. At the same time high band-width pathways needed for large throughput are situated in the intermediate network. Such abundant resources present in unwitting parts of the network are exploited by the attacker to launch a successful flooding attack.
- The handlers or the masters, which are compromised hosts with special programs running on them, are capable of controlling multiple agents.
- The attack daemon agents or zombie hosts are compromised hosts that are running a special program each and are responsible for generating a stream of packets towards the intended victim. These machines are commonly external to the victim’s own network to disable efficient response from the victim, and external to the network of the attacker to forswear liability if the attack is traced back.

2.1. DDoS Strategy

A Distributed Denial of Service (DDoS) attack is composed of several elements as shown in Figures 3 and 4.

There are several steps in launching a DDoS attack. These are shown in Figure 5.

1. Selection of agents. The attacker chooses the agents that will perform the attack. Based on the nature of vulnerabilities present, some machines are compromised to use as agents. Attackers victimize these machines, which have abundant resources, so that a powerful attack stream can be generated. In early years, the attackers attempted to acquire control of these machines manually. However, with the development of advanced security tool(s), it has become easier to identify these machines automatically and instantly.

2. Compromise. The attacker exploits security holes and vulnerabilities of the agent machines and plants the attack code. Not only that, the attacker also takes necessary steps to protect the planted code from identification and deactivation. As per the direct DDoS attack strategy, shown in Figure 3, the compromised nodes, i.e., zombies between the attacker and victim are recruited unwitting accomplice hosts from a large number of unprotected hosts connected through the Internet in high bandwidth. On the other hand, the DDoS attack strategy shown in Figure 4 is more complex due to inclusion of intermediate layer(s) between the zombies and victim(s). It further complicates the traceback mostly due to (i) complexity in untangling the traceback information (partial) with reference to multiple sources, and/or (ii) having to connect a large number of routers or servers. Self-propagating tools such as the Ramen worm [15] and Code Red [16] automate this phase. Unless a sophisticated defense mechanism is used, it is usually difficult for the users and owners of the agent systems to realize that they have become a part of a DDoS attack system. Another important feature of such an agent system is that the agent programs are very cost effective both in terms of memory and bandwidth. Hence they affect the performance of the system minimally.

3. Communication. The attacker communicates with any number of handlers to identify which agents are up and running, when to schedule attacks, or when to upgrade agents. Such communications among the attackers and handlers can be via various protocols, such as ICMP, TCP, or UDP. Based on configuration of the attack network, agents can communicate with a single handler or multiple handlers.

4. Attack. The attacker initiates the attack. The victim, the duration of the attack as well as special features of the attack such as the type, length, TTL, and port numbers can be adjusted. If there are substantial variations in the properties of attack packets, it is beneficial to the attacker, since it complicates detection.
In the past decade, attackers and agents have started using a multi-user, online chatting system known as Internet Relay Chat (IRC) [17]. This is because IRC chat networks allow users to create public, private and secret channels. An IRC-based DDoS attack network is similar to the Agent Handler DDoS attack model except that instead of using a handler program installed on a network server, an IRC server tracks the addresses of connected agents and handlers and facilitates communication among them. The discovery of the addresses of connected agents and handlers and facilitates communication among them. The discovery
of a single participant leads to the discovery of the communication channel, but other participants’ identities remain protected. Such multiuser online chatting systems or IRC systems have several other significant advantages for launching DDoS attacks. Among the three important benefits are: they afford a high degree of anonymity, they are difficult to detect, and they provide a strong, guaranteed delivery system. Furthermore, the attacker no longer needs to maintain a list of agents, since he can simply log on to the IRC server and see a list of all available agents [13]. IRC channels receive communications from the agent software regarding the status of the agents (i.e., up or down) and participate in notifying the attackers regarding the status of the agents. In an IRC-based DDoS attack, the agents are often referred to as “Zombie Bots” or “Bots”.

2.2. Generic architecture of DDoS attack defense mechanisms

Based on the locality of deployment, DDoS defense schemes can be divided into three classes: victim-end, source-end, and intermediate router defense mechanisms. All of these approaches have their own advantages and disadvantages. We discuss them one by one.

2.2.1. Victim-end defense mechanism

Victim-end detection approaches are generally employed in the routers of victim networks, i.e., networks providing critical Web services. A generic architecture of such schemes is shown in Figure 6. Here the detection engine is used to detect intrusion either online or offline, using either misuse based intrusion detection or anomaly based intrusion detection. The reference data stores information about known intrusion signatures or profiles of normal behavior. This information is updated by the processing elements as new knowledge about the observed behavior becomes available. The security manager often updates the stored intrusion signatures and also checks for other critical events such as false alarms. The processing element frequently stores intermediate results in the configuration data.

Detecting DDoS attacks in victim routers is relatively easy because of the high rate of resource consumption. It is also the most practically applicable type of defense scheme as Web servers providing critical services always try to secure their resources for legitimate users. But the problem with these approaches is that, during DDoS attacks, victim resources, e.g., network bandwidth, often gets overwhelmed and these approaches cannot stop the flow beyond victim routers. Another important disadvantage is that, these approaches detect the attack only after it reaches the victim and detecting an attack when legitimate clients have already been denied is not useful.

2.2.2. Source-end defense mechanism

A generic architecture of source-end preventive schemes is shown in Figure 7. This architecture is similar to the victim-end detection architecture. Here a throttling component is added to impose rate limit on outgoing connections. The observation engine compares both incoming and outgoing traffic statistics with some predefined normal profiles.

Detecting and stopping a DDoS attack at the source is the best possible defense. It prevents the possibility of flooding not only on the victim side, but also in the whole intermediate network. The main difficulty with this approach is that, detecting DDoS attacks at source end is not easy. This is because in these attacks, sources are widely distributed and a single source behaves almost similarly as in normal traffic. Another problem is the difficulty of deploying system at the source end.

2.2.3. Intermediate network defense mechanism

The intermediate network defense scheme balances the trade-offs between detection accuracy and attack bandwidth consumption, the main issues in source-end and victim-end detection approaches. Figure 8 shows a generic architecture of the intermediate network defense scheme, one that can be employed in any network router. Such a scheme is generally collaborative in nature and the routers share their observations with other routers. Like a source-end scheme, these schemes also impose rate limits on connections passing by the router after comparing with stored normal profiles.

Detection and traceback of attack sources are easy in this approach due to collaborative operation. Routers can form an overlay mesh to share their observations [18]. The main difficulty with this approach is deployability. To achieve full detection accuracy, all routers on the Internet will have to employ this detection scheme, because unavailability of this scheme in only a few routers may cause failure to the detection and traceback process. Obviously, full practical implementation of this scheme is extremely difficult by
reconfiguring all the routers on the Internet.

3. EXISTING METHODS FOR DDoS ATTACK DETECTION

In this section, we present a summary of existing literature on DDoS attack detection methods. These methods are based on the architectures discussed above namely, victim-end, source-end and in-network. We discuss these schemes without considering their practical deployability in real networks. Recent trends show that soft computing approaches have been used heavily for DDoS attack detection. Ensembles of classifiers have also performed satisfactorily with high detection rates. We classify methods for DDoS attack detection into four major classes as shown in Figure 9.

3.1. Statistical methods

Statistical properties of normal and attack patterns can be exploited for detection of DDoS attacks. Generally a statistical model for normal traffic is fitted and then a statistical inference test is applied to determine if a new instance belongs to this model. Instances that do not conform to the learnt model, based on the applied test statistics, are classified as anomalies. Chen et al. [19] develop a distributed change point (DCP) detection architecture using change aggregation trees (CATs). The non-parametric CUSUM approach was adapted to describe the distribution of pre-change or post-change network traffic. When a DDoS flooding attack is being launched, the cumulative deviation is noticeably higher than random fluctuations. The CAT mechanism is designed to work at the router level to detect abrupt changes in traffic flows. The domain server uses the traffic change patterns detected at attack-transit routers to construct the CATs, which represent the attack flow pattern. A very well-known DDoS defense scheme called D-WARD is presented in [20]. D-WARD identifies an attack based on continuous monitoring of bidirectional traffic flows between the network and the rest of the Internet and by periodic deviation analysis with the normal flow patterns. Mismatched flows are rate limited in proportion to their aggressiveness. D-WARD not only offers a good detection rate but also reduces DDoS attack traffic significantly. It uses a predefined model for normal traffic to detect anomalies.
in the two-way traffic statistics for each peer. If it identifies a DDoS attack, it imposes a rate limit on the suspicious outgoing flow for the peer. Next, D-WARD observes the traffic for either confirmation of the attack or refutation. If confirmed, D-WARD further controls the rate limit. However, if refuted, it gradually allows increased traffic rate. Saifullah [21] proposes a defense mechanism based on a distributed algorithm that performs weight-fair throttling at upstream routers. The throttling is weight-fair because the traffic destined for the server is controlled (increased or decreased) by leaky buckets at the routers based on the number of users connected, directly or through other routers, to each router. In the beginning of the algorithm, the survival capacity is underestimated by the routers so as to protect the server from any sudden initial attack. The rate is updated (increased or decreased), based on the server’s feedback sent to its child routers and eventually propagated downward to all routers, in the subsequent rounds of the algorithm with a view to converging the total server load to the tolerable capacity range.

Chen [22] presents a new detection method for DDoS attack traffic based on the two-sample t-test. It first obtains statistics for normal SYN arrival rate (SAR) and confirms that it follows the normal distribution. The method identifies an attack by computing (a) the difference between incoming SAR and normal SAR, and (b) the difference between the number of SYN and ACK packets. Unlike most previous DDoS defense schemes that only deal with either flooding or meek attack, the proposal uses two statistical tests to identify malicious traffic. It first compares the differences between the overall means of the incoming traffic arrival rate and the normal traffic arrival rate by the two-sample t-test. If the difference is significant, it concludes that the traffic may include flooding attack packets. However, the low-rate attack traffic may pass the arrival rate test and make the backlog queue full. The approach then compares the two groups that contain different numbers of SYN and ACK packets by the two-sample t-test. If there is a significant difference, it recognizes that the attack traffic is mixed into the current traffic. Zhang et al. [23] propose a prediction method for the available service rate of a protected server by applying the Auto Regressive Integrated Moving Average (ARIMA) model. They use available service rates to qualify the server’s availability to detect DDoS attacks. Their prediction method divides server resources into CPU time, memory utilization and networking buffer. Based on the prediction, they use abnormal detection technology to analyze the consumption of server resources to predict whether the server is under DDoS attack.

Akella et al. [24] explore key challenges in helping an ISP network detect attacks on itself or attacks on external sites which use the ISP network. They propose a detection mechanism where each router detects traffic anomalies using profiles of normal traffic constructed using stream sampling algorithms. Initial results show that it is possible to: (1) profile normal traffic reasonably accurately, (2) identify anomalies with low false positive and false negative rates (locally,
at the router), and (3) be cost-effective in terms of memory consumption and per packet computation. In addition, ISP routers exchange information with each other to increase confidence in their detection decisions. A router gathers responses from all other routers regarding suspicions and based on them decides whether a traffic aggregate is an attack or normal. The initial results show that individual router profiles capture key characteristics of the traffic effectively and identify anomalies with low false positive and false negative rates. Peng et al. [25] describe a novel approach to detect bandwidth attacks by monitoring the arrival rate of new source IP addresses. The detection scheme is based on an advanced non-parametric change detection scheme, CUSUM. Cheng et al. [26] propose the IP Flow Feature Value (FFV) algorithm based on the essential features of DDoS attacks, such as abrupt traffic change, flow dissymmetry, distributed source IP addresses and concentrated target IP addresses. Using a linear prediction technique, a simple and efficient ARMA prediction model is established for normal network flow. Then a DDoS attack detection scheme based on anomaly detection techniques and linear prediction model (DDAP) is used. Udhayan and Hamsapriya [27] present a Statistical Segregation Method (SSM), which samples the flow in consecutive intervals and compares the samples against the attack state condition and sorts them with the mean as the parameter. Then correlation analysis is performed to segregate attack flows from legitimate flows. The authors compare SSM against various other methods and identify a blend of segregation methods for alleviating false detections.

In [28], the authors introduce a generic DoS detection scheme based on maximum likelihood criterion with random neural networks (RNN). The method initially selects a set of traffic features in offline mode to obtain pdf estimates and to evaluate the likelihood ratios. During decision making, it measures the features of incoming traffic and attempts to decide according to each feature. Finally, it obtains an overall decision using both feed-forward and recurrent architectures of the RNN. A brief summary of these methods is given in Table 1

### 3.2. Soft computing methods

Learning paradigms, such as neural networks, radial basis functions and genetic algorithms are increasingly used in DDoS attack detection because of their ability to classify intelligently and automatically. Soft computing is a general term for describing a set of optimization and processing techniques that are tolerant of imprecision and uncertainty. Jalili et al. [29] introduce a DDoS attack detection system called SPUNNID based on a statistical pre-processor and unsupervised artificial neural nets. They use statistical pre-processing to extract features from the traffic, and an unsupervised neural net to analyze and classify traffic patterns as either a DDoS attack or normal.

Karimazad and Farahzad [30] propose an anomaly-based DDoS detection method based on features of attack packets, analyzing them using Radial Basis Function (RBF) neural networks. The method can be applied to edge routers of victim networks. Vectors with seven features are used to activate an RBF neural network at each time window. The RBF neural network is applied to classify data to normal and attack categories. If the incoming traffic is recognized as attack traffic, the source IP addresses of the attack packets are sent to the Filtering Module and the Attack Alarm Module for further actions. Otherwise, if the traffic is normal, it is sent to the destination. RBF neural network training can be performed as an off-line process but it is used in real time to detect attacks faster. Gavrilis and Dermatas [31] also present a detector for DDoS attacks in public networks based on statistical features estimated in short-time window analysis of incoming data packets. A small number of statistical descriptors are used to describe the behavior of the DDoS attacks. An accurate classification is achieved using Radial Basis Function neural networks. Wu et al. [32] propose to detect DDoS attacks using decision trees and grey relational analysis. The detection of the attack from the normal situation is viewed as a classification problem. They use 15 attributes, which not only monitor the incoming/outgoing packet/byte rate, but also compile the TCP, SYN, and ACK flag rates, to describe the traffic flow pattern. The decision tree technique is applied to develop a classifier to detect abnormal traffic flow. They also use a novel traffic pattern matching procedure to identify traffic flow similar to the attack flow and to trace back the origin of an attack based on this similarity. Nguyen and Choi [33] develop a method for proactive detection of DDoS attacks by classifying the network status. They break a DDoS attack into phases and select features based on an investigation of DDoS attacks. Finally, they apply the k-nearest neighbor (KNN) method to classify the network status in each phase of DDoS attack. A method presented in [34] detects DDoS attacks based on a fuzzy estimator using mean packet inter-arrival times. It detects the suspected host and traces the IP address to drop packets within 3 second detection windows.

Lately ensembles of classifiers have been used for DDoS attack detection. The use of an ensemble reduces the bias of existing individual classifiers. An ensemble of classifiers has been used by [35] for this purpose where a Resilient Back Propagation (RBP) neural network is chosen as the base classifier. The main focus of this paper is to improve the performance of the base classifier. The proposed classification algorithm, RBPBoost combines the output of the ensemble of classifier outputs and Neyman Pearson cost minimization strategy [36], for final classification decision. Table 2 presents a brief summary of the soft computing methods presented in this section.
3.3. Knowledge-based methods

In knowledge-based approaches, network events are checked against predefined rules or patterns of attack. In these approaches, general representations of known attacks are formulated to identify actual occurrences of attacks. Examples of knowledge-based approaches include expert systems, signature analysis, self organizing maps, and state transition analysis. Gil and Poletto [37] introduce a heuristic along with a data structure called MULTOPS (MULTi-Level Tree for Online Packet Statistics), that monitor certain traffic characteristics which can be used by network devices such as routers to detect and eliminate DDoS attacks. MULTOPS is a tree of nodes that contains packet rate statistics for subnet prefixes at different aggregation levels. Expansion and contraction of the tree occurs within a pre-specified memory size. A network device using MULTOPS detects ongoing bandwidth attacks by the presence of a significant and disproportional difference between packet rates going to and coming from the victim or the attacker. Depending on their setup and their location on the network, MULTOPS-equipped routers or network monitors may fail to detect a bandwidth attack that is mounted by attackers that randomizes IP source addresses on malicious packets. MULTOPS fails to detect attacks that deploy a large number of proportional flows to cripple a victim.

Thomas et al. [38] present an approach to DDoS...
defense called NetBouncer and claim it to be a practical approach with high performance. Their approach relies on distinguishing legitimate and illegitimate use and ensuring that resources are made available only for legitimate use. NetBouncer allows traffic to flow with reference to a long list of proven legitimate clients. If packets are received from a client (source) not on the legitimate list, a NetBouncer device proceeds to administer a variety of legitimacy tests to challenge the client to prove its legitimacy. If a client can pass these tests, it is added to the legitimacy list and subsequent packets from the client are accepted until a certain legitimacy window expires.

Wang et al. [39] present a formal and methodical way of modelling DDoS attacks using Augmented Attack Tree (AAT), and discuss an AAT-based attack detection algorithm. This model explicitly captures the particular subtle incidents triggered by a DDoS attack and the corresponding state transitions from the view of the network traffic transmission on the primary victim server. Two major contributions of this paper are: (1) an AAT-based DDoS model (ADDoSAT), developed to assess potential threat from malicious packets on the primary victim server and to facilitate the detection of such attacks; (2) an AAT-based bottom-up detection algorithm proposed to detect all kinds of attacks based on AAT modelling. Compared with the conventional attack tree modelling method, AAT is advanced because it provides additional information, especially about the state transition process. As a result, it overcomes the shortcomings of CAT modelling. There is currently no established AAT-based bottom up procedure for detecting network intrusions. Limwiwatkul and Rungsawang [40] propose to discover DDoS attack signatures by analyzing the TCP/IP packet header against well-defined rules and conditions, and distinguishing the difference between normal and abnormal traffic. The authors mainly focus on ICMP, TCP and UDP flooding attacks.

Zhang and Parashar [41] propose a distributed approach to defend against DDoS attacks by coordinating across the Internet. Unlike traditional IDS, it detects and stops DDoS attacks within the intermediate network. In the proposed approach, DDoS defence systems are deployed in the network to detect DDoS attacks independently. A gossip based communication mechanism is used to exchange information about network attacks between these independent detection nodes to aggregate information about the overall network attacks. Using the aggregated information, individual defence nodes obtain approximate information about global network attacks and can stop them more effectively and accurately. For faster and reliable dissemination of attack information, the network grows as a peer-to-peer overlay network on top of the Internet. Previously proposed approaches rely on monitoring the volume of traffic that is received by the victim. Most such approaches are incapable of differentiating a DDoS attack from a flash crowd. Lu et al. [42] describe a perimeter-based anti-DDoS system, in which the traffic is analyzed only at the edge routers of an Internet Service Provider (ISP) network. The anti-DDoS system consists of two major components: (1) temporal-correlation based feature extraction and (2) spatial-correlation based detection. The scheme can accurately detect DDoS attacks and identify attack packets without modifying existing IP forwarding mechanisms at routers. A brief summary of these knowledge based methods is given in Table 3.

### 3.4. Other data mining and machine learning methods

An effective defense system to protect network servers, network routers, and client hosts from becoming handlers, zombies, and victims of DDoS flood attacks is presented in [43]. The NetShield system protects any IP-based public network on the Internet. It uses preventive and deterrent controls to remove system vulnerabilities on target machines. Adaptation techniques are used to launch protocol anomaly detection and provide corrective intrusion responses. The NetShield system enforces dynamic security policies. NetShield is especially tailored for protecting network resources against DDoS flood attacks. Chen et al. [44] present a comprehensive framework for DDoS attack detection known as DDoS Container. It uses a network based detection method to overcome complex and evasive types of DDoS attacks. It works in inline mode to inspect and manipulate ongoing traffic in real time. By continuous monitoring of both DDoS attacks and legitimate applications, DDoS Container covers stateful inspection on data streams and correlates events among different sessions. It proactively terminates the session when it detects an attack. Lee et al. [45] propose a method for proactive detection of DDoS attacks by exploiting an architecture consisting of a selection of handlers and agents that communicate, compromise and attack. The method performs cluster analysis. The authors experiment with the DARPA 2000 Intrusion Detection Scenario Specific Dataset to evaluate the method. The results show that each phase of the attack scenario is partitioned well and can detect precursors of a DDoS attack as well as the attack itself. Sekar et al. [46] investigate the design space for in-network DDoS detection and propose a triggered, multi-stage approach that addresses both scalability and accuracy. Their contribution is the design and implementation of LADS (Large-scale Automated DDoS detection System). The system makes effective use of the data (such as NetFlow and SNMP feeds from routers) readily available to an ISP. Rahmani et al. [47] discuss a joint entropy analysis of multiple traffic distributions for DDoS attack detection. They observe that the time series of IP flow numbers and aggregate traffic sizes are strongly statistically dependant. The occurrence of an attack...
affects this dependence and causes a rupture in the time series for joint entropy values. Experimental results show that this method could lead to more accurate and effective DDoS detection.

A low-rate DDoS attack has significant ability to conceal its traffic because of its similarity with normal traffic. Xiang et al. [48] propose two new information metrics: (i) generalized entropy metric and (ii) information distance metric, to detect low-rate DDoS attacks. They identify the attack by measuring the distance between legitimate traffic and attack traffic. The generalized entropy metric is more effective than the traditional Shannon metric [49]. In addition, the information distance metric outperforms the popular Kullback-Leibler divergence approach. Francois et al. [50] present a method called FireCol based on information theory for early detection of flooding DDoS attacks. FireCol is comprised of an intrusion prevention system (IPS) located at the Internet service provider (ISP) level. The IPSs form virtual protection rings around the hosts to defend and collaborate by exchanging selected traffic information. The approach reported in [51] analyzes DDoS and flash crowd characteristics and provides an effective way to distinguish between the two in VoIP networks. The authors validate the method by simulation. A wavelet transformation and probability theory based network anomaly detection approach is proposed in [52]. The approach is able to identify known as well as unknown attacks. Zhong and Yue [53] present a DDoS attack detection model that extracts a network traffic model and a network packet protocol status model and sets the threshold for the detection model. Captured network traffic values are clustered based on the k-means clustering algorithm to build initial threshold values for network traffic. All captured packets are used to build the packet protocol status model using the Apriori [54] and FCM [55] algorithms. Whenever the current network traffic is over the threshold value, the network packet protocol status is checked to detect abnormal packets. If there are no abnormal packets, the current network traffic is clustered again by the k-means module to build a new threshold value model.

A two-stage automated system is proposed in [56] to detect DoS attacks in network traffic. It combines the traditional change point detection approach with a novel one based on continuous wavelet transforms [57]. The authors test the system using a set of publicly available attack-free traffic traces superimposed with anomaly profiles. In [58], Li and Lee present a systematic wavelet based method for DDoS attack detection. They use energy distribution based on wavelet analysis to detect DDoS attack traffic. Energy distribution over time has limited variation if the traffic keeps its behavior over time. Gupta et al. [59] use ANN to estimate the number of zombies in a DDoS attack. They use sample data to train a feed-forward neural network generated using the NS-2 network simulator. The generalization capacity of the trained network is promising and the network is able to predict the number of zombies involved in a DDoS attack with test error. A port-to-port specific traffic in a router, called IF flow is introduced in [60]. An important feature of IF is that it can amplify the attack to normal traffic ratio. An RLS (recursive least square) filter is used to predict IF flows. Next, a statistical method using a residual filtered process is used to detect anomalies. Finally, the authors applied the method to three types of traffic: IF flows, input links and output links within a router, and compare the anomaly detection results using ROC curves. Results show that IF flows are more powerful than input links and output links for DDoS attack detection. Cheng et al. [61] propose the IAI (IP Address Interaction Feature) algorithm considering interactions among addresses, abrupt traffic changes, many-to-one asymmetries among addresses, distributed source IP addresses and concentrated target addresses [61]. The IAI algorithm is designed to describe the essential characteristics of network flow states. Furthermore, a support vector machine (SVM) classifier, which is trained by an IAI time series from normal flow and attack flow, is applied to classify the

### TABLE 3. Knowledge based DDoS attack detection methods

<table>
<thead>
<tr>
<th>Reference</th>
<th>Objective</th>
<th>Deployment</th>
<th>Mode of working</th>
<th>Trainable</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gil and Polesello [37]</td>
<td>Attack prevention</td>
<td>Between source and destination network</td>
<td>Centralized</td>
<td>No</td>
<td>Each network devices maintains a data structure known as MULTOPS. Fails to detect attacks that deploy a large number of DDoS attack flows using a large number of agents, IP spoofing attacks.</td>
</tr>
<tr>
<td>Thomas et al. [38]</td>
<td>Attack detection</td>
<td>Victim side</td>
<td>Centralized</td>
<td>No</td>
<td>Network-aware interneate DDoS traffic from flash crowd using inline packet processing based on network processor technology.</td>
</tr>
<tr>
<td>Linowatkan and Rungrusawang [40]</td>
<td>Attack detection</td>
<td>Victim side</td>
<td>Distributed</td>
<td>Yes</td>
<td>Uses a TCP packet header to construct attack signature model for DDoS attack detection.</td>
</tr>
<tr>
<td>Zhang and Parashar [41]</td>
<td>Proactive</td>
<td>Intermediate network</td>
<td>Distributed</td>
<td>Yes</td>
<td>A gossip based scheme uses to get global information about DDoS attacks by information sharing.</td>
</tr>
<tr>
<td>Lu et al. [42]</td>
<td>Attack detection</td>
<td>Edge router</td>
<td>Distributed</td>
<td>Yes</td>
<td>Exploits spatial and temporal correlation of DDoS attack traffic records for detecting anomalous packet.</td>
</tr>
<tr>
<td>Wang et al. [39]</td>
<td>Attack detection</td>
<td>Victim side</td>
<td>Centralized</td>
<td>No</td>
<td>Uses an Augmented Attack Tree model for the detection of DDoS attacks and also can detect other attacks.</td>
</tr>
</tbody>
</table>
state of current network flows and identify the DDoS attacks. Experimental results show that the IAI-based detection scheme can distinguish between normal flows and abnormal flows with DDoS attacks effectively, and help identify fast and accurate attack flows when the attacking traffic is hidden among a relatively large volume of normal flows or close to the attacking sources. In addition, it has higher detection and lower false alarm rates compared to competing techniques.

The method presented in [62] can identify flooding attacks in real time and also can assess the intensity of the attackers based on fuzzy reasoning. The process consists of two stages: (i) statistical analysis of the network traffic time series using discrete wavelet transform and Schwarz information criterion (SIC) to find the change point of the Hurst parameters resulting from DDoS flood attack, and then (ii) identification and assessment of the intensity of the DDoS attack adaptively based on an intelligent fuzzy reasoning mechanism. Test results by ns2 based simulation with various network traffic characteristics and attack intensities demonstrate that the method could detect DDoS flood attack timely, effectively and intelligently. Zhang et al. [63] present a CPR (Congestion Participation Rate) based approach to detect low-rate DDoS (LDDoS) attacks using flow level network traffic. A flow with higher CPR value leads to LDDoS and consequent dropping of the packets. The authors evaluate the mechanism using ns2 simulation, testbed experiments and Internet traffic trace and claim that the method can detect LDDoS flows effectively. Another protocol specific feature based DDoS attack detection mechanism is introduced in [5]. It identifies a most relevant subset of features using correlation and can detect DDoS attacks with high detection accuracy.

In [64], a mathematical model is presented to provide gross evaluation of the benefits of DDoS defence based on dropping of attack traffic. Simulation results and testbed experiments are used to validate the model. In the same work, the authors also consider an autonomic defence mechanism based on CPN (Cognitive Packet Network) protocol and establish it to be capable of tracing back flows coming into a node automatically. Ghanea-Hercock et al. [65] provide a survey of the techniques within the Hyperion project. They also suggest an overall system architecture to improve the situational awareness of field commanders by providing an option to fuse and compose information services in real time. In [66], Gelenbe describes an approach to develop a self-aware networks to end user the option to explore the state of the network to find the best ways to meet their communication needs. In [67], a model is introduced for searching by $N$ agents in an unbounded random environment. The model allows for the loss or destruction of searchers and finite lifetime. A summarized presentation of these methods in this category is given in Table 4.

3.5. Discussion

Exact comparison of DDoS attack detection schemes is not feasible because some papers do not specify their results clearly, whereas others evaluate their schemes using different datasets or in different testing conditions. A comparison (as shown in Table 5) establishes that most papers do not consider all the issues that are pertinent. For example, the Distributed Change Point detection method [19] performs well for TCP SYN attacks, but its performance degrades for UDP attacks with large packet sizes. D-WARD [20] fails to detect pulsing attacks, especially when the inactive period is large. For NetBouncer [38], the legitimacy tests may not be exhaustive and certain illegitimate clients may also pass the test. In addition, Netbouncer is overwhelmed by flash crowds. Moreover, the delay introduced by the test affects new legitimate clients. The DDoS Detection system based on K-means and FCM clustering [53] performs well for unknown attacks, but the trade-off between detection accuracy and speed is high. Detection using RBF Neural Networks and statistical features [31] performs well for known attacks, but no dynamic modification can be performed easily for unknown attacks.

4. PERFORMANCE EVALUATION

Performance evaluation is important for any DDoS attack defense system. Performance evaluation is highly dependent on (i) the approach (ii) deployment status and (iii) whether there is facility for dynamic updation of profiles. While designing a DDoS attack defense scheme, these parameters should be taken into consideration. In this section, we primarily discuss the datasets that have been used for evaluating performance of detection methods. We also briefly introduce DDoS tools.

4.1. DDoS tools

There are many tools available to launch DDoS attacks in the literature [68, 44]. The architectures are almost always the same. Some are made by the attackers with slightly modifying others. Table 6 presents some of the tools with brief descriptions.

4.2. Datasets

After a new or enhanced detection mechanism is developed, it needs to be validated with proper datasets before deployment in a real life network. This is because evaluation in real traffic is extremely difficult or not possible because of non-availability of capturing tools from a high speed network. The proper choice of network intrusion datasets plays a vital role during evaluation of any computer attack detection method. First, the dataset should contain correctly captured real network traffic. Second, it should be unbiased. Only
The Computer Journal, Vol. ??, No. ??, ???

TABLE 4. Data mining and machine learning based DDoS attack detection methods

<table>
<thead>
<tr>
<th>Reference</th>
<th>Objective</th>
<th>Deployment</th>
<th>Mode of working</th>
<th>Trainable</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Li and Lee [58]</td>
<td>Attack detection</td>
<td>Victim end</td>
<td>Centralized</td>
<td>No</td>
<td>An energy distribution based wavelet analysis technique for the detection of DDoS traffic.</td>
</tr>
<tr>
<td>Sekar et al. [46]</td>
<td>Attack detection</td>
<td>Source side</td>
<td>Distributed</td>
<td>Yes</td>
<td>A triggered multi-stage approach for both scalability and accuracy for DDoS attack detection.</td>
</tr>
<tr>
<td>Gelenbe and Loukas [64]</td>
<td>Attacks defense using packet dropping</td>
<td>Victim end</td>
<td>Centralized</td>
<td>Yes</td>
<td>Detects attack by tracing back flows automatically.</td>
</tr>
<tr>
<td>Lee et al. [45]</td>
<td>Attack detection</td>
<td>Source side</td>
<td>Centralized</td>
<td>Yes</td>
<td>Detects DDoS attack proactively based on cluster analysis with agent handler architecture.</td>
</tr>
<tr>
<td>Rahinani et al. [47]</td>
<td>Attack detection</td>
<td>Victim side</td>
<td>Distributed</td>
<td>No</td>
<td>A joint entropy analysis of multiple traffic distributions for DDoS attack detection.</td>
</tr>
<tr>
<td>Li and Li [52]</td>
<td>Detectors of DDoS attacks automatically</td>
<td>Victim end</td>
<td>Centralized</td>
<td>Yes</td>
<td>A DDoS attack detection model based on wavelet transformation and probability theory.</td>
</tr>
<tr>
<td>Damotti et al. [56]</td>
<td>Detection of DDoS attack anomalies</td>
<td>Victim end</td>
<td>Centralized</td>
<td>Yes</td>
<td>Detects attacks correctly using combination of traditional change point detection and continuous wavelet transformation.</td>
</tr>
<tr>
<td>Xia et al. [62]</td>
<td>Detects flood attack and its intensity</td>
<td>Victim end</td>
<td>Centralized</td>
<td>Yes</td>
<td>A method to detect DDoS flooding attack using fuzzy logic.</td>
</tr>
<tr>
<td>Xiang et al. [48]</td>
<td>Detects low rate flooding attacks</td>
<td>Victim end</td>
<td>Centralized</td>
<td>No</td>
<td>Detects low-rate DDoS flooding attacks using new information flooding metrics effectively.</td>
</tr>
<tr>
<td>Gupta et al. [59]</td>
<td>Number of zombies identification</td>
<td>Victim end</td>
<td>Distributed</td>
<td>Yes</td>
<td>Uses ANN to estimate the number of zombies in a DDoS attack.</td>
</tr>
<tr>
<td>Franco et al. [50]</td>
<td>DDoS flooding attack detection</td>
<td>Source end</td>
<td>Distributed</td>
<td>Yes</td>
<td>A complete DDoS flooding attack detection technique. Also support incremental deployment in real network.</td>
</tr>
</tbody>
</table>

(i) Benchmark datasets: Benchmark datasets are prepared using an enterprise network and with proper labels such as normal or attack. Only a few benchmark intrusion datasets are publicly available but they are not for DDoS attacks. The KDDCup99 intrusion dataset is the most popular intrusion dataset publicly available but it is not suitable for DDoS attacks.

(ii) Simulated datasets: Another alternative for evaluating DDoS attack detection methods is to simulate the environment using available tools such as ns2, QualNet, and OMNeT+++. These simulators generally generate traffic statistics randomly and sometimes use statistical distributions. But it is difficult to ensure that the generated traffic correlates well with real network traffic.

(iii) Private datasets: The best approach for testing any intrusion detection system or DDoS attack detection method is to create a real network testbed with a large number of host and network components. Then one can generate normal as well as attack traffic and capture them using standard tools, used for high speed networks. After capture, one preprocesses the raw data and extracts features in a distributed manner and finally correlates each feature to create a complete dataset. After necessary labelling, it should be useful to test any intrusion detection system.

5. OPEN ISSUES AND CHALLENGES

Many methods for DDoS detection have been reported in the literature, but only a few of them have been applied in a real network environment and work effectively. Designing and implementing an ideal and practical DDoS defense system is really difficult. The main challenges that any DDoS defense scheme must overcome to become ideally usable are presented below.

(i) In a DDoS attack, attackers try to make a service unavailable to its legitimate clients and launch the attack using a large number of zombies distributed in different networks. It takes only a few seconds to exhaust the bandwidth and other resources of the victim. A faster detection scheme usually consumes higher processing power and at the same time, it adversely affects detection accuracy. An
TABLE 5. A general comparison of DDoS Attack Detection methods. Column 4 in the table represents if the method can be used real-time (R) or non-real time (N). The Detection rate (DR) and False Alarm Rate (FAR) are given in column 8 and column 9, respectively.

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Approach</th>
<th>Architecture/Method</th>
<th>R/N</th>
<th>Scalability</th>
<th>Unknown attack detection</th>
<th>Dynamic signature updation</th>
<th>DR with dataset used</th>
<th>FAR with dataset used</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCD approach [19]</td>
<td>statistical</td>
<td>DCP / Change aggregation tree</td>
<td>R</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>98% (flooding attacks)</td>
<td>&lt; 1% (flooding attacks)</td>
</tr>
<tr>
<td>Weighted fair model [21]</td>
<td>statistical</td>
<td>Weight Fair Throttling</td>
<td>R</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SPUNNID model [29]</td>
<td>statistical</td>
<td>SPUNNID system / unsupervised neural network</td>
<td>R</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>94.9% (flooding attacks)</td>
<td>5% (flooding attacks)</td>
</tr>
<tr>
<td>D-WARD system [20]</td>
<td>knowledge based</td>
<td>Self regulating reverse feedback system / rate limited</td>
<td>R</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>(flooding attacks)</td>
<td>0.5% (flooding attacks)</td>
</tr>
<tr>
<td>MULTIOPS system [37]</td>
<td>knowledge based</td>
<td>Multi-level tree based method</td>
<td>N</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>(IP spoofing attacks)</td>
<td>-</td>
</tr>
<tr>
<td>NetBouncer model [38]</td>
<td>knowledge based</td>
<td>Packet filtering method</td>
<td>R</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>(real-time data)</td>
<td>-</td>
</tr>
<tr>
<td>Decision tree model [32]</td>
<td>knowledge based</td>
<td>Decision tree based method</td>
<td>R</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>98% (flooding attacks)</td>
<td>2.4% (flooding attacks)</td>
</tr>
<tr>
<td>Clustering model [53]</td>
<td>statistical</td>
<td>Data mining based method</td>
<td>N</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>98.65% (real-time)</td>
<td>1.12% (real-time)</td>
</tr>
<tr>
<td>RBF neural net model [30]</td>
<td>soft computing</td>
<td>RBF system / Radial basis function</td>
<td>R</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>98.2% (UCLA data)</td>
<td>0.01% (UCLA data)</td>
</tr>
<tr>
<td>RBF neural net model [31]</td>
<td>soft computing</td>
<td>Radial-basis-function based method</td>
<td>R</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>98%-100% (real-time)</td>
<td>0% (real-time)</td>
</tr>
<tr>
<td>T-test model [22]</td>
<td>statistical</td>
<td>T-test method</td>
<td>R</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>98%-100% (flooding attacks)</td>
<td>5%-7% (flooding attacks)</td>
</tr>
<tr>
<td>ARMA based model [23]</td>
<td>statistical</td>
<td>Prediction based method</td>
<td>R</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>(new simulation)</td>
<td>1%-3% (new simulation)</td>
</tr>
<tr>
<td>Attack tree model [19]</td>
<td>knowledge based</td>
<td>Attack tree based method</td>
<td>R</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>(flooding attacks)</td>
<td>-</td>
</tr>
<tr>
<td>Signature discovery approach [40]</td>
<td>knowledge based</td>
<td>Traffic statistics based method</td>
<td>R</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>(flooding attacks)</td>
<td>-</td>
</tr>
<tr>
<td>Profile based approach [24]</td>
<td>statistical</td>
<td>Stream sampling based method</td>
<td>R</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>(IP spoofing attacks)</td>
<td>2% (IP spoofing attacks)</td>
</tr>
<tr>
<td>Cooperative model [41]</td>
<td>knowledge based</td>
<td>RL-DDoS system / Gossip-based scheme</td>
<td>R</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>(Emulab simulation)</td>
<td>7%-12% (Emulab simulation)</td>
</tr>
<tr>
<td>Sequential non-parametric change point method [25]</td>
<td>knowledge based</td>
<td>Nonparametric CPD method</td>
<td>R</td>
<td>yes</td>
<td>yes</td>
<td>No</td>
<td>90%-100% (Auckland traces)</td>
<td>-</td>
</tr>
<tr>
<td>Perimeter based anti-DDoS system [42]</td>
<td>knowledge based</td>
<td>Spatial correlation based method</td>
<td>R</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>93.0% (IP spoofing attacks)</td>
<td>0.05% (IP spoofing attacks)</td>
</tr>
<tr>
<td>R-NN classifier approach [33]</td>
<td>statistical</td>
<td>Nearest neighbour based method</td>
<td>R</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>91.88% (DARPA 2000 DDoS data)</td>
<td>8.11% (DARPA 2000 DDoS data)</td>
</tr>
<tr>
<td>Change point detect by fuzzy logic [62]</td>
<td>Soft computing</td>
<td>Fuzzy logic based method</td>
<td>R</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>(ns2 simulation)</td>
<td>-</td>
</tr>
<tr>
<td>Linear prediction model [26]</td>
<td>statistical</td>
<td>Linear prediction based method</td>
<td>R</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>96.1% (DDoS flow data, LL-DDoS 2.0.2)</td>
<td>0.8% (DDoS flow data, LL-DDoS 2.0.2)</td>
</tr>
<tr>
<td>SSM method [27]</td>
<td>statistical</td>
<td>Statistical segregation based method</td>
<td>R</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>(CAIDA data)</td>
<td>1.2% (CAIDA data)</td>
</tr>
<tr>
<td>Ensemble of neural net model [35]</td>
<td>soft computing</td>
<td>Ensemble of neural net based classifiers</td>
<td>N</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>99.4% (DARPA 2000 DDoS data)</td>
<td>5.7% (DARPA 2000 DDoS data)</td>
</tr>
<tr>
<td>Autonomic mathematical model [64]</td>
<td>Machine learning</td>
<td>CPN based method</td>
<td>R</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>(ns2 simulation)</td>
<td>-</td>
</tr>
</tbody>
</table>

An offline classification scheme generally ensures good performance as classification can be performed in a more sophisticated manner in comparison to real-time detection. But accurately detecting all attacks after interrupting services to legitimate clients has no use. So, emphasis should be given more to speed over accuracy of detection for a practical DDoS defense mechanism.

(ii) Real-time detection of low-rate DDoS attack with high detection accuracy and low false alarm is
a challenging task, since such traffic follows the normal traffic distribution.

(iii) A DDoS defense mechanism cannot simply be judged based on its performance with a standard fixed dataset containing normal and a few attack

<table>
<thead>
<tr>
<th>Name and Ref.</th>
<th>Description</th>
<th>Protocol</th>
<th>Attack</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trinoo [69, 70]</td>
<td>• Widely used by the attackers as well as research community.</td>
<td>UDP</td>
<td>UDP flood</td>
</tr>
<tr>
<td></td>
<td>• A bandwidth depletion attack tool, used to launch coordinated UDP flood attacks against one or many IP addresses.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Fixed size UDP packets are sent to the victim machine's random ports.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Does not spoof source addresses.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Implements UDP Flood attacks against the target victim.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tribe Flood Network (TFN) [71]</td>
<td>• Able to wage both bandwidth depletion and resource depletion attacks.</td>
<td>UDP, ICMP, TCP</td>
<td>TCP, SYN flood, ICMP flood, smurf</td>
</tr>
<tr>
<td></td>
<td>• Uses a command line interface to communicate between the attacker and the control master program.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Offers no encryption between agents and handlers or between handlers and the attacker.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Allows TCP SYN and ICMP flood as well as smurf attacks.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TFN2K [72]</td>
<td>• Developed using the TFN DDoS attack tool.</td>
<td>TCP, UDP, ICMP</td>
<td>smurf, SYN flood, UDP flood, ICMP flood</td>
</tr>
<tr>
<td></td>
<td>• Adds encrypted messaging among all of the attack components [73].</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Communications between real attacker and control master program are encrypted using a key-based CAST-256 algorithm [74].</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Conducts covert exercises to hide itself from intrusion detection systems.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Can forge packets that appear to come from neighboring machines.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Provides other options such as TARGA and MIX attack [75].</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stacheldraht [76]</td>
<td>• Based on early versions of TFN and eliminates some of its weak points by combining features of Trinoo.</td>
<td>TCP, UDP, ICMP</td>
<td>TCP, SYN flood, UDP flood, ICMP flood echo request flood</td>
</tr>
<tr>
<td></td>
<td>• Performs updates on the agents automatically.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Provides a secure telnet connection via symmetric key encryption among the attackers and handlers.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Communicates through TCP and ICMP packets.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mstream [77]</td>
<td>• Uses spoofed TCP packets with the ACK flag set to attack the target.</td>
<td>TCP, UDP</td>
<td>TCP ACK flood</td>
</tr>
<tr>
<td></td>
<td>• A simple point-to-point TCP ACK flooding tool to overwhelm the tables used by fast routing routines in switches.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Communications are not encrypted, and performed through TCP/UDP packets; zombie is connected via telnet by master.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Target gets hit by ACK packets and sends TCP RST to non-existent IP addresses.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Routers return “ICMP unreachable” causing more bandwidth starvation.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Possesses very limited control features and can spoof by randomizing all 32 bits of the source IP address.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shaft [78]</td>
<td>• A successor of Trinoo.</td>
<td>TCP, UDP, ICMP</td>
<td>TCP/UDP/ICMP flood</td>
</tr>
<tr>
<td></td>
<td>• Uses UDP communication between handlers and agents.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Shaft provides UDP/ICMP/TCP flooding attack options; it randomizes source IP address and source port in packets.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• The size of packets remains fixed during the attack.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Able to switch the handler’s IP address and port in real time during the attack.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Able to switch control master servers and ports in real time, hence making detection by intrusion detection tools difficult.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trinity v3 [79]</td>
<td>• Various TCP floods are used by randomizing all 32-bits of the source IP address, such as TCP fragment floods, TCP established floods, TCP RST packet floods, and TCP random flag packet floods.</td>
<td>TCP, UDP</td>
<td>TCP fragment floods, TCP RST packet floods, TCP random flag packet floods, TCP established floods</td>
</tr>
<tr>
<td></td>
<td>• Generates TCP flood packets with random control flags set to provide a wider set of TCP based attacks.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knight [80]</td>
<td>• A very lightweight yet powerful IRC based attack tool.</td>
<td>TCP, UDP</td>
<td>UDP flood, SYN and PUSH+ACH flood</td>
</tr>
<tr>
<td></td>
<td>• Provides SYN attacks, UDP Flood attacks, and an urgent pointer flooder [81].</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Designed to run on Windows operating systems and has features such as an automatic updater via http or ftp, a checksum generator and more.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Uses Trojan horse program called Back Orifice for installation in the target host.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LORC [8]</td>
<td>• A powerful anonymous attacking tool via IRC.</td>
<td>TCP, UDP, HTTP</td>
<td>UDP, TCP, HTTP flood</td>
</tr>
<tr>
<td></td>
<td>• Operates in three methods of attack: TCP, UDP and HTTP.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Exists in two versions: binary version and web-based version.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The Computer Journal, Vol. ??, No. ??, ???
packets. It must be scalable to real networks for actual deployment. In a real application, a defense scheme should be able to capture, process and classify incoming packets in real time. It cannot just expect to store captured packets and analyze them later, obviously increasing the time lag between the occurrence of an event and its subsequent detection. Particularly, in the case of DDoS attacks, the problem would become more acute, as all resources will be exhausted. So all phases of detection should run in a parallel manner in real time. Generally, real time DDoS detection systems are expected to be scalable for use in high speed real networks.

(iv) With the continuing progress in Internet technologies, attackers are developing and launching new attacks with greater sophistication day by day. So, detection of new attacks is a challenge for any defense system. In supervised learning, the classifier needs to be trained with a proper minimum training dataset, not biased towards any attack. This means that it cannot detect unknown attacks, particularly those with behavior very different from existing attack classes. Unsupervised approaches are appropriate for detecting unknown attacks. However, these methods are usually not suitable from real time performance. Hence, developing a combined approach based on both supervised and unsupervised approaches with the capability of detecting both known and unknown attacks real time or near real time is of utmost necessity.

(v) Accurate segregation of high-rate DDoS attack traffic from normal flash crowds with minimum resource consumption or low false alarm rate in real-time or near real-time is a challenging task.

(vi) Transparency to existing Internet infrastructure is very important in terms of deployment. It should not be allowed to slow down the processing of normal packets from legitimate clients, which itself is denial of service. Moreover, the defense system itself should not be vulnerable to attacks, creating unavoidable breakdown of service. Most importantly, a DDoS defense scheme should be deployable in real networks.

(vii) High speed traffic analysis for detecting DDoS attacks is a challenging task. A defense scheme can get overwhelmed in real networks, particularly at the time of a DDoS attack. So, the defense scheme should be able to handle the crowd and still function properly. A defense scheme capable of real time detection should perform well with high speed traffic. Offline schemes suffer in high speed traffic because of the overhead created by processing delay, which can slow down detection speed further or can cause total breakdown in extreme cases.

(viii) Real time updation of network statistics and fast identification of randomized spoofed IP addresses are challenges.

(ix) In DDoS attacks with a large number of agents, attack behavior often conforms well with normal behavior. In such a situation, for a DDoS defense mechanism aiming to provide a near real time solution may have to be based on an incremental clustering algorithm to segregate the attack from normal traffic. This requires an appropriate proximity measure that works sensibly, quickly and reliably.

(x) The detection method should be dependent on a minimum number of input parameters if not independent of parameters and should also be based on a minimum number of traffic parameters or features.

6. POSSIBLE SOLUTIONS AGAINST DDoS ATTACKS

DDoS defense schemes are of three types based on their deployment: source-end, victim-end and intermediate router defense mechanisms. Among these three, most researchers are in favor of using the victim-end approach. However, a common disadvantage of this scheme is the consumption of a huge amount of resources to provide fast detection response. In the latest DDoS attacks 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 they may extend activities to complex database query transactions also. Generally, segregation of such transactions from the legitimate queries is a difficult task.

Current DDoS attacking tools are capable of launching attacks in different modes [20] such as increasing rate, constant rate, pulsing (attack rate oscillates between maximum to 0) and gradual pulsing (e.g., attack rate achieves a maximum in 20 seconds and reduces to 0 in 10 seconds). These various forms of attacks may involve single as well as multiple attack sources, that too using single as well as multiple source addresses. Defenders are interested in providing a semi-automatic solution with an objective of achieving reduced false alarm with minimum resource consumption. A solution we propose at a high level is based on an ensemble approach working in a distributed framework using appropriate fusion techniques for making decision. Protocol specific feature extraction in a distributed environment involving multiple sensors and detecting DDoS attacks based on a class-specific subset of features at the individual level will provide the base for the underlying layer of the proposed solution. Finally, the individual decisions can be combined using an appropriate combination rule at the upper layer.
at any of the participating nodes in the distributed framework for the final inference. Once the attack occurrence is confirmed, the next task is to limit the attack rate instantly in a selective manner without affecting service to legitimate users, so that subsequent damage can be minimized immediately. However, for significant reduction of the false alarm rate, involvement of human experts is often useful. Especially, in the case of low-rate DDoS attacks, an appropriate diagnosis of false alarms with the help of human analysts, and then providing feedback to the detecting sensors, are very essential. The three major advantages of our solution are: (i) it helps achieve an unbiased and high detection rate at reduced false alarms, (ii) it minimizes resource consumption by sharing the computation cost and (iii) it achieves a scalable, real-time detection performance.

7. CONCLUSION

While developing a DDoS defense scheme, the issues discussed in this paper need to be deliberated and considered with due seriousness. In this paper, we have presented an overview of DDoS attacks, detection schemes and finally research issues and challenges. In addition, we provide a comparison among current detection methods. Practically designing and implementing a DDoS defense is very difficult. The comparison of the existing detection mechanisms shows that most schemes are not capable of fulfilling all the requirements for real-time network defense. Different performance parameters need to be balanced against each other delicately and appropriately. A possible solution to counter DDoS attacks is also briefly outlined in this paper.

ACKNOWLEDGMENT

This work is supported by the Department of Information Technology and the Council of Scientific & Industrial Research (CSIR), Government of India. The authors are thankful to the funding agencies.

REFERENCES

Ad Hoc and Ubiquitous Computing.

and grey relational analysis.

Networks and ISDN Systems.

RBF networks and statistical features.


Based method for DDoS attacks detection using rbf networks and statistical features.


