Detection and Prevention of DDOS Attacks Using Group Testing Approach

Mr. Y. Sathishkanna , Dr.S.Pathurnisha

Abstract— Asymmetric application layer DDoS attacks using computationally intensive HTTP requests are an extremely dangerous class of attacks capable of taking down web servers with relatively few attacking connections. These attacks consume limited network bandwidth and are similar to legitimate traffic, which makes their detection difficult. Existing detection mechanisms for these attacks use indirect representations of actual user behaviour and complex modelling techniques, which leads to a higher false positive rate (FPR) and longer detection time, which makes them unsuitable for real time use. There is a need for simple, efficient and adaptable detection mechanisms for asymmetric DDoS attacks. In this work, an attempt is made to model the actual behavioral dynamics of legitimate users using a simple annotated Probabilistic Timed Automata (PTA) along with a suspicion scoring mechanism for differentiating between legitimate and malicious users. This allows the detection mechanism to be extremely fast and have a low FPR. In addition, the model can incrementally learn from run-time traces, which makes it adaptable and reduces the FPR further. Experiments on public datasets reveal that our proposed approach has a high detection rate and low FPR and adds negligible overhead to the web server, which makes it ideal for real time use.

Keywords— detection mechanisms, Probabilistic Timed Automata (PTA), malicious users, negligible overhead etc.

I. Introduction

Distributed Denial of Service (DDoS) attacks using computationally intensive HTTP requests have emerged as a serious threat to web applications in recent years. These attacks are called asymmetric Application Layer DDoS(AL-DDoS) attacks and are capable of exhausting server resources using considerably fewer attack requests than other application layer DDoS attacks.

In addition to their potency, Asymmetric DDoS attacks possess certain features which make them extremely difficult to detect. First, they are executed using legitimate HTTP requests, which makes it impossible to detect these attacks by inspecting individual requests. Second, they exhibit a very low attack bandwidth and thus, cannot be detected by existing volumetric DDoS detection mechanisms. Third, they resemble legitimate user traffic such as flash crowds, which are sudden spikes in legitimate user traffic to a web server due to a noteworthy event or sale. Our contributions in this work are as follows:

- We propose the use of an annotated Probabilistic Timed Automata (PTA) to capture the behavioral dynamics of legitimate users accessing a web application.
- We propose a mechanism to detect asymmetric AL-DDoS attacks by using cumulative suspicion score assignment based on the annotated PTA.
- We demonstrate that the proposed detection mechanism performs considerably well in detecting asymmetrical-DDoS attacks, and can be used effectively at real time.

II. LITERATURE REVIEW

[1] J. Francois, I. Aib, and R. Boutaba, “Firecol: a Collaborative Protection Network for the Detection of Flooding ddos Attacks,”

Distributed Denial Of Service (DDoS) attacks remain a major security problem, the mitigation of which is very hard especially when it comes to highly distributed botnet-based attacks. The early discovery of these attacks, although challenging, is necessary to protect end-users as well as the expensive network infrastructure resources. In this Project, we address the problem of DDoS attacks and present the theoretical foundation, architecture, and algorithms of FireCol. The core of FireCol is composed of intrusion prevention systems (IPSs) located at the Internet service providers (ISPs) level. The IPSs form virtual protection rings around the hosts to defend and collaborate by exchanging selected traffic information. The evaluation of FireCol using extensive simulations and a real dataset is presented, showing FireCol effectiveness and low overhead, as well as its support for incremental deployment in real networks.

2. S. T. Zargar, J. Joshi, D. Tipper, “A Survey of Defense Mechanisms Against Distributed Denial Of Service (DDoS) Flooding Attacks,”

Distributed Denial Of Service (DDoS) flooding attacks are one of the biggest concerns for security professionals. DDoS flooding attacks are typically explicit attempts to disrupt legitimate users' access to services. Attackers usually gain access to a large number of computers by exploiting their vulnerabilities to set up attack armies (i.e., Botnets). Once an attack army has been set up, an attacker can invoke a coordinated, large-scale attack against one or more targets. Developing a comprehensive defense mechanism against identified and anticipated DDoS flooding attacks is a desired goal of the intrusion detection and prevention research

Mr. Y. Sathishkanna , PG Scholar, Department of Computer Science and Engineering , Nehru Institute of Technology, Coimbatore -641105, (Email Id: sathishkannacse@yahoo.com)

Dr.S.Pathurnisha, Professor and HOD, Department of Computer Science and Engineering , Nehru Institute of Technology, Coimbatore -641105, (Email Id: nitpathurnisha@nehrucolleges.com)

community. However, the development of such a mechanism requires a comprehensive understanding of the problem and the techniques that have been used thus far in preventing, detecting, and responding to various DDoS flooding attacks. In this Project, we explore the scope of the DDoS flooding attack problem and attempts to combat it. We categorize the DDoS flooding attacks and classify existing countermeasures based on where and when they prevent, detect, and respond to the DDoS flooding attacks. Moreover, we highlight the need for a comprehensive distributed and collaborative defense approach. Our primary intention for this work is to stimulate the research community into developing creative, effective, efficient, and comprehensive prevention, detection, and response mechanisms that address the DDoS flooding problem before, during and after an actual attack.

3. A. Yaar, A. Perrig, D. Song, “StackPi: New Packet Marking and Filtering Mechanisms for DDoS and IP Spoofing Defense,”

Today's Internet hosts are threatened by large-scale distributed Denial Of Service (DDoS) attacks. The path identification (Pi) DDoS defense scheme has recently been proposed as a deterministic packet marking scheme that allows a DDoS victim to filter out attack packets on a per packet basis with high accuracy after only a few attack packets are received (Yaar , 2003). In this Project, we propose the StackPi marking, a new packet marking scheme based on Pi, and new filtering mechanisms. The StackPi marking scheme consists of two new marking methods that substantially improve Pi's incremental deployment performance: Stack-based marking and write-ahead marking. Our scheme almost completely eliminates the effect of a few legacy routers on a path, and performs 2-4 times better than the original Pi scheme in a sparse deployment of Pi-enabled routers. For the filtering mechanism, we derive an optimal threshold strategy for filtering with the Pi marking. We also develop a new filter, the PiIP filter, which can be used to detect Internet protocol (IP) spoofing attacks with just a single attack packet. Finally, we discuss in detail StackPi's compatibility with IP fragmentation, applicability in an IPv6 environment, and several other important issues relating to potential deployment of StackPi.

4. H. Wang, C. Jin, K. G. Shin, “Defense Against Spoofed IP Traffic Using Hop-Count Filtering,”

IP spoofing has often been exploited by Distributed Denial Of Service (DDoS) attacks to: 1)conceal flooding sources and dilute localities in flooding traffic, and 2)coax legitimate hosts into becoming reflectors, redirecting and amplifying flooding traffic. Thus, the ability to filter spoofed IP packets near victim servers is essential to their own protection and prevention of becoming involuntary DoS reflectors. Although an attacker can forge any field in the IP header, he cannot falsify the number of hops an IP packet takes to reach its destination. More importantly, since the hop-count values are diverse, an attacker cannot randomly spoof IP addresses while maintaining consistent hop-counts. On the other hand, an Internet server can easily infer the hop-count information from the Time-to-Live (TTL) field of the IP header. Using a mapping between IP addresses and their hop-counts, the server can distinguish spoofed IP packets from legitimate ones. Based on this observation, we present a novel filtering technique, called Hop-Count Filtering (HCF)-which builds an accurate IP-to-hop-count (IP2HC) mapping table-to detect and discard spoofed IP packets. HCF is easy to deploy, as it does not require any support from the underlying network. Through analysis using network measurement data, we show that HCF can identify close to 90% of spoofed IP packets, and then discard them with little collateral damage. We implement and evaluate HCF in the Linux kernel, demonstrating its effectiveness with experimental measurements.

III. EXISTING SYSTEM:

The demonstrate how lack of strong location authentication allows creation of software-based Sybil devices that expose crowd sourced map systems to a variety of security and privacy attacks.

This experiments show that a single Sybil device with limited resources can cause havoc on Waze, reporting false congestion and accidents and automatically rerouting user traffic.

More importantly, the describe techniques to generate Sybil devices at scale, creating armies of virtual vehicles capable of remotely tracking precise movements for large user populations while avoiding detection.

A. DISADVANTAGES:

□ Less security and privacy
□ Highly impact of the attacks

IV. PROPOSED SYSTEM:

The propose a new approach based on co-location edges, authenticated records that attest to the one-time physical co location of a pair of devices. Over time, co-location edges combine to form large proximity graphs that attest to physical interactions between devices, allowing scalable detection of virtual vehicles.

The demonstrate the efficacy of this approach using large-scale simulations, and how they can be used to dramatically reduce the impact of the attacks. They have informed Waze/Google team of our research findings. Currently, they are in active collaboration with Waze team to improve the security and privacy of their system.

A. ADVANTAGES

□ Improve security and privacy
□ Reduce impact of the attacks
□ Large proximity graphs
□ This approach using large-scale simulations

V.SYSTEM ARCHITECTURE

A. SYSTEM DEVELOPMENT MODULES

- Login Process Denial of Services.

- Group attacker modules.
- Group testing modules.
- Victim/Detection modules

VI. MODULE DESCRIPTION

A. LOGIN PROCESS DENIAL OF SERVICES

It may be possible to overwhelm the login process by continually sending login-requests that require the presentation tier to access the authentication mechanism, rendering it unavailable or unreasonably slow to respond.

When a user enters an incorrect username and/or password, the application should respond with a generic error message stating that the information entered was incorrect. If the application explicitly states which component of the username/password pair was incorrect then an attacker can automate the process of trying common usernames from a dictionary file in an attempt to enumerate the users of the application. Whilst applications may handle authentication failure messages correctly, many still allow attackers to enumerate users through the forgotten password feature.

B. GROUP ATTACKER MODULES.

The maximum destruction caused by the attacks includes the depletion of the application service resource at the server side, the unavailability of service access to legitimate user, and possible fatal system errors which require rebooting the server for recovery. We assume that any malicious behaviors can be discovered by monitoring the service resource usage, based on dynamic value thresholds over the monitored objects. Data manipulation and system intrusion are out of this scope. That application interface presented by the servers can be readily discovered and clients communicate with the servers using HTTP/1.1 sessions on TCP connections.

We consider a case that each client provides a non spoofed ID, which is utilized to identify the client during our detection period. Despite that the application DDoS attack is difficult to be traced; by identifying the IDs of attackers the firewall can block the subsequent malicious requests. The attackers are assumed to launch application service requests either at high inter arrival rate or high workload, or even both. The term “request” refers to either main request or embedded request for HTTP page. Since the detection scheme proposed will be orthogonal to the session affinity, we do not consider the repeated one-shot attack mentioned in. We further assume that the number of attackers d << n where n is the total client amount. This arises from the characteristics of this attack. Due to the benefits of virtual server s we employee, this constraint can be relaxed, but we keep it for the theoretical analysis in the current work.

C. GROUP TESTING MODULES

The classic GT model consists of t pools and n items (including at most d positive ones). This model can be represented by a t _ n binary matrix M where rows represent the pools and columns represent the items. An entry M[I, j]= 1 if and only if the I th pool contains the j th item; otherwise, M[I, j]= 0. The t-dimensional binary column vector V denotes the test outcomes of these t pools, where 1-entry represents a positive outcome and 0-entry represents a negative one. Note that a positive outcome indicates that at least one positive item exists within this pool; whereas negative one means that all the items in the current pool are negative.

A detection model based on GT can be assume that there are t virtual servers and n clients, among which d clients are Binary testing matrix M and testing outcome vector V. Attackers. Consider the matrix M t*n in Fig. 1, the clients can be mapped into the columns and virtual servers into rows in M, where M[I, j]= 1 if and only if the requests from client j are distributed to virtual server i. With regard to the test outcome column V, we have V[i]= 1 if and only if virtual server i has received malicious requests from at least one attacker, but we cannot identify the attackers at once unless this virtual server is handling only one client. Otherwise, if V ½i_ ¼ 0, all the clients assigned to server I are legitimate. The d attackers can then be captured by decoding the test outcome vector V and the matrix M.

D. VICTIM/DETECTION MODULES

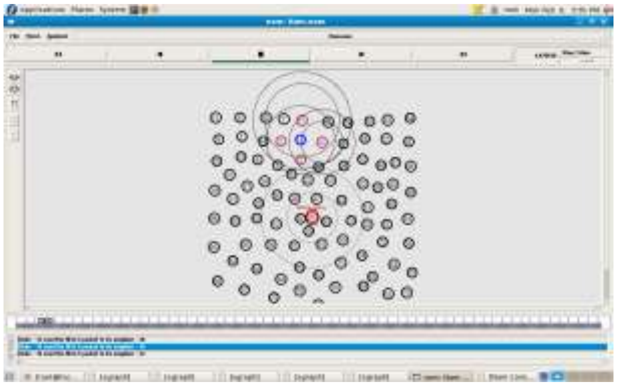
The victim model in our general framework consists of multiple back-end servers, which can be Web/application servers, database servers, and distributed file systems. We do not take classic multitier Web servers as the model, since our detection scheme is deployed directly on the victim tier and identifies the attacks targeting at the same victim tier; thus, multitier attacks should be separated into several classes to utilize this detection scheme. We assume that all the back-end servers provide multiple types of application services to clients using HTTP/1.1 protocol on TCP connections.

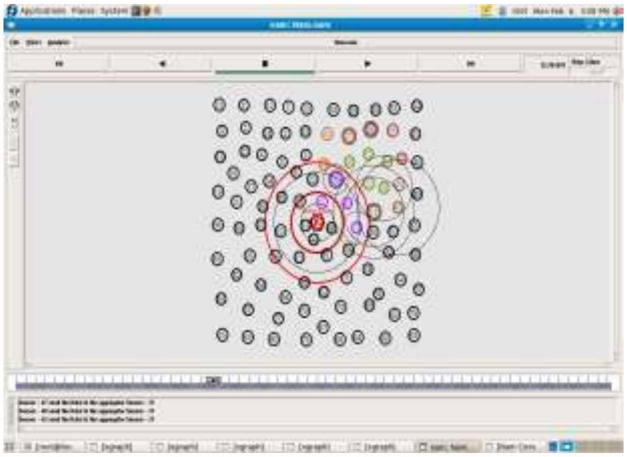
Each back-end server is assumed to have the same amount of resource. Moreover, the application services to clients are provided by K virtual private servers (K is an input parameter), which are embedded in the physical back-end server machine and operating in parallel. Each virtual server is assigned with equal amount of static service resources, e.g., CPU, storage, memory, and network bandwidth. The operation of any virtual server will not affect the other virtual servers in the same physical machine. There a sons for utilizing virtual servers are twofold: first, each virtual server can reboot independently, thus is feasible for recovery from possible fatal destruction; second, the state transfer overhead for moving clients among different virtual servers is much smaller than the transfer among physical server machines.

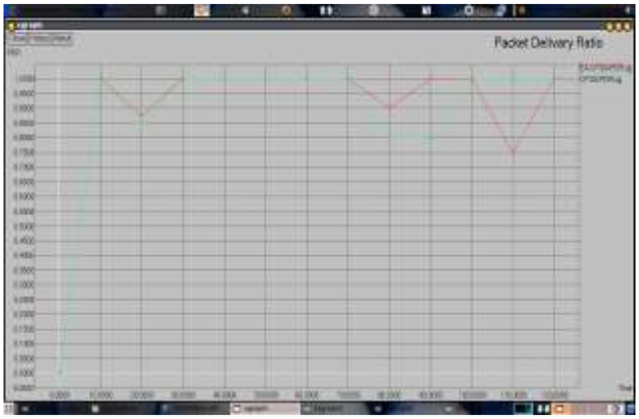
VII. RESULTS AND DISCUSSION

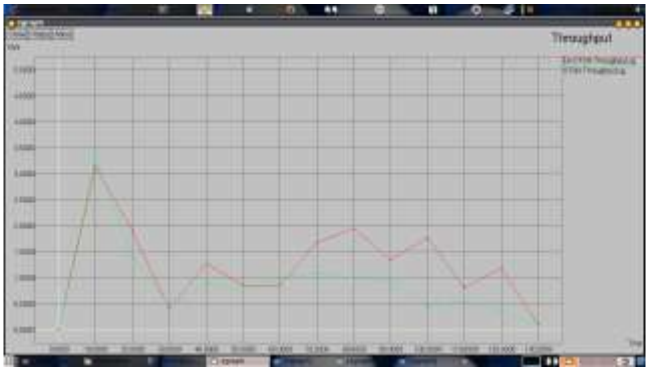
Once the network layout is generated, traffic source is to be added to the both legitimate user and attackers. The attacker traffic source is formed in such a way that it won’t obey the traffic rules and try to utilize the full network resources. If any network congestion takes place, it won’t slowdown its traffic generation rate and it keeps on sending packets. The legitimate user follows the traffic rules and if any congestion takes place, it slowdowns its sending rate. After the congestion, it checks whether the route is free or not. If the route is still congested, it waits for some more time the network traffic to clear. The legitimate user requests the server and the server opens a connection for it and the legitimate user uses the connection like three way hand shake. Attacker requests the server then the server opens the connections, but the attacker won't use the connection and keeps on requesting many connections. Due to this the server connections are wasted and sever is not able to serve legitimate user.

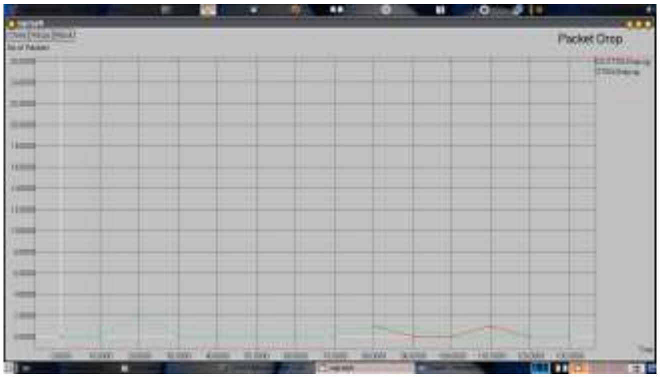
Pushback Mechanism Phase This is the more important phase, where the intelligent router initiates pushback mechanism and there to client puzzles. In this phase, a router in the ISP network is selected as intelligent router and a powerful mechanism is implemented on this router.











The implementation of pushback mechanism is as follows. Making congestion Signature, Matching the traffic pattern, updating Congestion Signature and Initiation of Pushback mechanism. With the proposed method, attack traffic over the network is greatly reduced. The proposed system effectively provides defense against DoS/DDoS attacks.

The effect of defense mechanism and decrease in the attacker traffic over the network. This method identifies the attacker hosts before the attack. Figure 6 shows the amount of attacker traffic over the network before the proposed method and the amount of attacker traffic after the defense mechanism. In this method suspected host is validated with client Puzzles and after the conformation attacker traffic is dropped at the edge routers. Figure 7 shows the legitimate user traffic strength over the network before and after the proposed system implementation. The attacker traffic consumes much network resources before it reaches the victim (target) thus leading to congestion. With the proposed system, such problems are rectified as the congestion signature can be adjusted and updated in order to find these abnormalities over the network.

VIII.CONCLUSION

A novel technique for detecting application DDOS attack by means of a new constraint-based group testing model. Motivated by classic GT methods, three detection algorithms were proposed and a system based on these algorithms was introduced. Theoretical analysis and preliminary simulation results demonstrated the outstanding performance of this system in terms of low detection latency and false positive/negative rate.

Our focus of this Project is to apply group testing principles to application DDOS attacks, and provide an underlying framework for the detection against a general case of network assaults, where malicious requests are indistinguishable from normal ones. For the future work, we will continue to investigate the potentials of this scheme and improve this proposed system to enhance the detection efficiency. Some possible directions for this can be:

1. The sequential algorithm can be adjusted to avoid the requirement of isolating attackers

2. More efficient d-disjunct matrix could dramatically decrease the detection latency, as we showed in the theoretical analysis. A new construction method for this is to beproposedand can be a major theoretical work for anotherProject.

3. The overhead of maintaining the state transfer among virtual servers can be further

decreased by more sophisticated techniques.

4. Even that we already have quite low false positive/ negative rate from the algorithms.

We can still improve it via false-tolerant group testing methods. This error-tolerant matrix has great potentials to improve the performance of the PND algorithm and handle application DDOS attacks more efficiently.

IX. FUTURE ENHANCEMENT

We will continue to investigate the potentials of this scheme and improve this proposed system to enhance the detection efficiency.

□ The sequential algorithm can be adjusted to avoid the requirement of isolating attackers.

□ More efficient d-disjunct matrix could dramatically decrease the detection latency, as we showed in the theoretical analysis. A new construction method for this is to be proposed and can be a major theoretical work for another Project.

□ The overhead of maintaining the state transfer among virtual servers can be further decreased by more sophisticated techniques.

□ Even that we already have quite low false positive/ negative rate from the algorithms, we can still improve it via false-tolerant group testing methods.

REFERENCES

[1] J. Francois, I. Aib, and R. Boutaba, "Firecol: a Collaborative Protection Network for the Detection of Flooding ddos Attacks," IEEE/ACM Trans. on Netw., vol. 20, no. 6, Dec. 2012, pp. 1828-1841.

[2] S. T. Zargar, J. Joshi, D. Tipper, "A Survey of Defense Mechanisms Against Distributed Denial Of Service (DDoS) Flooding Attacks," IEEE Commun. Surv. & Tut., vol. 15, no. 4, pp. 2046 - 2069, Nov. 2013.

[3] A. Yaar, A. Perrig, D. Song, "StackPi: New Packet Marking and Filtering Mechanisms for DDoS and IP Spoofing Defense," IEEE J. on Sel. Areas in Commun., vol. 24, no. 10, pp. 1853 - 1863, Oct. 2006.

[4] H. Wang, C. Jin, K. G. Shin, "Defense Against Spoofed IP Traffic Using Hop-Count Filtering," IEEE/ACM Trans. on Netw., vol. 15, no. 1, pp. 40 - 53, Feb. 2007.

[5] Z. Duan, X. Yuan, J. Chandrashekar, "Controlling IP Spoofing through Interdomain Packet Filters," IEEE Trans. on Depend. and Secure Computing, vol. 5, no. 1, pp. 22 - 36, Feb. 2008.

[6] M. Sung, J. Xu, "IP traceback-based intelligent packet filtering: a novel technique for defending against Internet DDoS attacks," IEEE Trans. On Parall. and Distr. Sys., vol. 14, no. 9, pp. 861 - 872, Sep. 2003.

[7] M. Sung, J. Xu, J. Li, L. Li, "Large-Scale IP Traceback in High-Speed Internet: Practical Techniques and Information-Theoretic Foundation," IEEE/ACM Trans. on Netw., vol. 16, no. 6, pp. 1253 - 1266, Dec. 2008.

[8] Y. Xiang, K. Li, W. Zhou, "Low-Rate DDoS Attacks Detection and Traceback by Using New Information Metrics," IEEE Trans. on Inf. Foren. and Sec., vol. 6, no. 2, pp. 426 - 437, May 2011.

[9] H. Ballani, Y. Chawathe, S. Ratnasamy, T. Roscoe, S. Shenker, "Off by default!," In Proc. HotNets-IV, Nov. 2005, College Park, MD, USA.

[16] A. Yaar, A. Perrig, and D. Song, "SIFF: a stateless internet flow filter to mitigate DDoS flooding attacks," In Proc. IEEE Symposium on Security and Privacy, May 2004, Oakland, CA, USA.

[10] H. Luo, Z. Chen, J. Cui, H. Zhang, M. Zukerman, C. Qiao, "CoLoR: an information-centric internet architecture for innovations," IEEE Network, vol. 28, no. 3, pp. 4 - 10, May 2014.