Research article Available online www.ijsrr.org ISSN: 2279–0543

International Journal of Scientific Research and Reviews

Anomaly Detection on IP Flow Using Bivariate Parametric Detection Mechanism

K.V. Lakshmi Praveena * and K. Brahma Reddy

M.Tech. Scholar, CSE, NBKRIST, Vidyanagar, Spsr Nellore, India

ABSTRACT

Security in computer networks is an extremely active and broad area of research, as networks of all sizes are targeted daily by attackers seeking to disrupt or disable network traffic. A successful denial-of-service attack degrades network performance, resulting in losses of several millions of dollars. Development of methods to counter these and other threats is thus of high interest. Current countermeasures under development focus on detection of anomalies and intrusions, their prevention, or a combination of both.

In this paper, we present an anomaly detection method on IP flow by using bivariate parametric detection mechanism (bPDM) which profiles normal traffic; a traffic-rate shift and a change in the distribution of packet-sizes from the nominal condition is flagged as an anomaly. Our anomaly detection problem is posed as a statistical hypothesis test. We develop parametric statistical models for typical and anomalous traffic. The detection method does not need, or attempt, to model the full traffic patterns; instead it captures key, gross features of the traffic to enable informed decisions about changes in traffic. We underscore that our model does not capture all aspects of general Internet traffic. This model effectively captures changes in the traffic which are associated with network anomalies. Our goal is to see whether these simple, approximate statistical models can yield detection methods of high performance by modeling sufficient, salient features of the traffic.

KEYWORDS: Anomaly, Denial of services (DoS), Sequential Probability Ratio Test (SPRT), bPDM Algorithm.

***Corresponding Author**
K.V. Lakshmi Praveena
M.Tech. Scholar (CSE)
Dept. of. CSE
NBKRIST, Vidyanagar
SPSR Nellore, AP
E – mail- Praveena1506@gmail.com
Mob. No. - 9032326165

INTRODUCTION

The Rule based anomaly detection on IP flow deals with detection of anomalies in the aggregate traffic. The proposed bivariate parametric detection mechanism (bPDM) uses a sequential probability ratio test, allowing for control over the false positive rate while examining the tradeoff between detection time and the strength of an anomaly. Additionally, it uses both traffic-rate and packet-size statistics, yielding a bivariate model that eliminates most false positives. The method is analyzed using the bit-rate signal-to-noise ratio (SNR) metric, which is shown to be an effective metric for anomaly detection. The performance of the bPDM is evaluated in three ways. First, synthetically generated traffic provides for a controlled comparison of detection time as a function of the anomalous level of traffic. Second, the approach is shown to be able to detect controlled artificial attacks over the University of Southern California (USC), Los Angeles, campus network in varying real traffic mixes. Third, the proposed algorithm achieves rapid detection of real denial-of-service attacks as determined by the replay of previously captured network traces.[1]

EXISTING SYSTEM

The existing system present an anomaly detection method that profiles normal traffic; a traffic-rate shift and a change in the distribution of packet sizes from the nominal condition is flagged as an anomaly. The evaluated methods using synthetic traces and emulated Iperf attacks, and find that the bPDM can detect attacks in a few seconds. The detection times for the synthetic attacks are validated using real and proxy-real network attacks, and the bit-rate SNR is shown to be not only an effective metric for evaluating anomaly detection methods, but also a better one than the previously proposed packet SNR metric. The anomaly detection is developed only for the limited number of hops i.e.., in the example the number of hops between two areas is limited to 8-10 hops.[2]

PROPOSED SYSTEM

The proposed system deals with increase or decrease in the number of hops. Here the hop count is increased or decreased and tested using the tool called “Iperf”. This Iperf tool is the testing tool that can create TCP and UDP data streams and measure the throughput of the network. The increases in the hops are tested using the “Iperf” tool and checked that there is no effect in the anomaly detection. So the proposed system tries to prove that increase in the hop count will not affect the anomaly detection.

TYPES OF ATTACKS

In terms of the number of malicious entities involved in an attack, we distinguish: uni source attacks – launched by and originating from a single source;

Distributed Attacks

Distributed Attacks are launched by and Originating from multiple coordinated sources, though not necessarily involving more than one malicious end user. Distributed DoS attacks operate on a much broader scale (with practically limitless number of launch sites) and can considerably add to the severity, length and scale of an attack, making it possible to practically disable even very powerful servers over prolonged periods of time.[3] Such was the case with the servers of large commercial sites like Yahoo!, eBay.com, etc. in early February 2000. Since then, distributed attacks have turned from a theoretical possibility to a major concern for Internet servers of any size and computing power.

Uni Source Attacks

Uni Source Attacks are launched by and originating from a single source. Many modern operating systems incorporate interrupt-driven network subsystem architectures, which have been shown to lack both efficiency and stability under conditions of high network load. The problem comes from the fact that they give strictly highest priority to processing of incoming network packets, regardless of which application those packets belong to, whether or not this application is currently executing and whether or not this receiver application has lower priority that the currently executing one. As a result, a situation known as receiver livelock could potentially occur, where the network server spends all of its resources processing incoming packets, only to later discard them because no CPU time was left to service application programs.[4] Denial of service attacks could disable servers for potentially long periods of time. During that time between the onset of such an attack and the time when the breach is actually detected and recovered from, the victim server is unable to handle any requests by legitimate non-malicious users. For large commercial servers this translates to a significant loss of income, and which they consider even more serious – a loss of reputation. To be concrete, one report estimated the total loss due to the distributed attacks we mentioned in the range of $1.2 billion.

NETWORK ANOMALIES

Network anomalies typically refer to circumstances when network operations deviate from normal network behavior. Network anomalies can arise due to various causes such as malfunctioning network devices, network overload, malicious denial of service attacks, and network intrusions that

disrupt the normal delivery of network services.[5]These anomalous events will disrupt the normal behavior of some measurable network data. In this paper, we present techniques that can be employed to detect such types of anomalies. The definition of normal network behavior for measured network data is dependent on several network specific factors such as the dynamics of the network being studied in terms of traffic volume, the type of network data available, and types of applications running on the network. Accurate modeling of normal network behavior is still an active field of research, especially the online modeling of network traffic.

PROPOSED TECHNOLOGY

The proposed system deals with increase or decrease in the number of hops. Here the hop count is increased or decreased and tested using the tool called "Iperf". This Iperf tool is the testing tool that can create TCP and UDP data streams and measure the throughput of the network. The increase in the hops is tested using the "Iperf" tool and checked that there is no effect in the anomaly detection. So the proposed system proves that increase in the hop count will not affect the anomaly detection.

Anomaly Detection Using Parametric Model

The SPRTs for the packet-rate and packet-size features that are the primary components of the bPDM. The bPDM operates on a unidirectional sampled time-series of aggregate network traffic. The parametric models employed to derive the bPDM are not representative of general Internet traffic, but rather are chosen to differentiate between the presence-of-anomaly and background-only hypotheses.

A classical SPRT assumes known and constant model parameters. In reality, such parameter values are not always available, and thus we consider a generalized likelihood ratio test (GLRT), defined as

$$G_N(\mathbf{x}) = \prod_{k=1}^{N} \frac{p(x_k, \widehat{\Theta}_1 | H_1)}{p(x_k, \widehat{\Theta}_0 | H_0)}$$

where we use the conditional probability density with their maximum likelihood(ML) estimates. To form the generalized SPRT, the estimated parameters are substituted into the test form as previously described. In particular, the model parameters are updated using non overlapping windows. We initially use fixed-size windows for both hypotheses; a 1-s sliding window ensures that enough data is being collected to derive good estimates of the background and attack parameters, denoted s. The offset window employed to estimate the parameters uses more recent samples, and thus the change in the model parameters can be detected as evidenced in Section V. Whenever the SPRT crosses the lower

threshold, confirming the absence of an attack, the ASN function is computed under hypothesis , and the update window size is reset to

$$M = \min\{\mathbb{E}_0(N), M_{\text{init}}\}.$$

Similarly, when an attack is detected by the bPDM, the length of the update window for the parameters is reset to where the first argument of the min functions in and are the ASN functions under hypotheses H_0 and H_1, respectively, and have been derived in .We now derive the SPRTs for both the packet-rate and packet-size features, and then describe the bPDM algorithm.

$$N = \min\{\mathbb{E}_1(N), N_{\text{init}}\}$$

Detecting SPRT for the Packet Rate

The null hypothesis H_0, which represents only background traffic, is modeled using the generalized Poisson distribution (GPD), whose probability density function (pdf) is given by

$$p(x|H_0) = \theta(\theta + \lambda x)^{x-1} e^{-\theta-\lambda x}/x!$$

Where $x \in \{0, 1, \ldots\}$ is the number of packet arrivals in a fixed time interval and $\{\theta, \lambda\}$ are the parameters of the GPD. We model an anomaly or attack stream as a constant-rate source with deterministic, unknown rate. Our work focuses on detecting a set of commonly occurring attacks, which is a class of attacks such as DoS attacks that use fixed-size attack packets. Since DoS attacks are also characterized by the attacker flooding the network, this set of attacks corresponds to the constant-rate attack traffic assumption made above. However, as evidenced in Section V-E, the bPDM can also quickly and accurately detect smart attacks which employing varying packet sizes. A random variable drawn from the anomalous distribution is specified as

$$Y = r + X$$

Where is drawn from the GPD distribution that models the background-only hypothesis. For the anomaly hypothesis, we assume that the constant-rate anomaly follows the pdf of the shifted GPD (sGPD)1 given by

$$p(x|H_1) = \theta\,(\theta + \lambda(x-r))^{x-r-1}\, e^{-\theta-\lambda(x-r)}/(x-r)!$$

The SPRT, in the case of the packet-rate feature, requires us to compare the generalized likelihood ratio

$$G_N(\mathbf{x}) = \prod_{k=1}^{N} \frac{p(x_k, \hat{\theta}_1, \hat{\lambda}_1, \hat{r}|H_1)}{p(x_k, \hat{\theta}_0, \hat{\lambda}_0|H_0)}$$

to the threshold given in (3). Note that the densities specified in (10) are the GPD (7) and sGPD (9) with parameter estimates used in lieu of known parameter values.

Incorporating the Packet-Size SPRT

The packet-size distribution of normal Internet traffic has been characterized in [27] as mostly bimodal at 40 and 1500 bytes (with 40% and 20% of packets, respectively). An examination of our background trace data, which include Ethernet and VLAN headers, validates this model but with differing means. The background traffic in our traces can also be characterized as mostly bimodal, with means at 68 and 1518 bytes, which represent approximately 40% and 20% of the total packets, respectively. We note, however, that no specific distribution is ascertained for the remaining 40% of the packets. We expect packet-size distribution information to be effective in attack detection since a broad class of attacks use a single packet size; e.g., DNS reflector attacks use the maximum packet size, and TCP SYN attacks use the minimum packet size. Thus, the influx of attack packets, in the case of attacks that employ a single attack packet size, will alter the relative number of a specific packet size with respect to the packet-size distribution of normal traffic.[6] As such, the sample entropy of the packet-size distribution can be used to distinguish between the background only and presence-of-anomaly hypotheses. In the bPDM framework, recall that represents the number of packet arrivals in the interval $[i/p, (i+1)/p)$. Let $\mathbb{S}_i$ denote the set of distinct packet sizes that arrive in this interval, and denote the proportion of packets of size to the total number of packets in the same interval. Thus, the sample entropy is computed as

$$y_i = -\sum_{j \in \mathbb{S}_i} q_j \log q_j .$$

The sample entropy is modeled using the Gaussian distribution given by

$$p(y|H_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{1}{2\sigma_i^2}(y-\mu_i)^2\right]$$

for both the background and attack hypotheses. Thus, the log-likelihood ratio (LLR), given observations, is specified as

$$\log L(\mathbf{y}) = a_2 \sum_{i=1}^{N} y_i^2 + a_1 \sum_{i=1}^{N} y_i + a_0$$

The resulting SPRT requires that we continue to take more observations if

$$\log(A) < \log G(\mathbf{y}) < \log(B)$$

CONCLUSION

The existing system present an anomaly detection method that profiles normal traffic; a traffic-rate shift and a change in the distribution of packet sizes from the nominal condition is flagged as an anomaly. The small change in the packet size or packet rate is flagged as the anomaly detection. This method is true when there are no any disturbances in the traffic. The anomaly detection is developed only for the limited number of hops i.e.., in the example the number of hops between two areas is limited to 8-10 hops.

The proposed system deals with increase or decrease in the number of hops. Here the hop count is increased or decreased and tested using the tool called “Iperf”. This Iperf tool is the testing tool that can create TCP and UDP data streams and measure the throughput of the network. The increase in the hops is tested using the “Iperf” tool and checked that there is no effect in the anomaly detection. So the proposed system proves that increase in the hop count will not affect the anomaly detection. Although this conjecture is consistent with the data, its verification in a controlled experiment is an area for future work.

REFERENCES

1. Gautam Thatte, Urbashi Mitra and John Heideman, Senior Member, IEEE, ACM. “Parametric Methods for Anomaly Detection in Aggregate Traffic.” Ming Hseih Dept. of Electr. Eng., Univ. of Southern California, Los Angeles, CA, USA. April. 2011; 9: 319-24.
2. L. Feinstein, D. Schnackenberg, R. Balupari and D. Kindred. “Statistical approaches to DDoS attack detection and response,” in Proc.DARPA Inf. Survivability Conf. Expos., 2003, 2(4): 303–314.
3. J. Ellis and T. Speed, The Internet Security Guidebook: From Planning to Deployment. New York: Academic, 2001; 1 (1): 30–32.

4. George Nychis, Vyas Sekar, David G. Andersen, Hyong Kim, Hui Zhang "An Empirical Evaluation of Entropy-based Traffic Anomaly Detection" 2007; 4 (4) 3–4.
5. Valentin Razmov "Denial of Service Attacks and How to Defend Against Them" May 10, 2000; 41(1): 10–12.
6. Marina Thottan and Chuanyi Ji, "Anomaly Detection in IP Networks SIGNALPROCESSING, AUGUST 2003; 51(8): 10-11.