A Prediction-based Detection Algorithm against Distributed Denial-of-Service Attacks

Guoxing Zhang
South China University of Technology
Guangdong Nortel Tel. Eq. Co. Ltd.
Guangzhou, China

Shengming Jiang (Senior member, IEEE)
University of Glamorgan, UK
South China University of Technology
Guangzhou, China
Email: jiangsm@ieee.org

Gang Wei (Member, IEEE)
South China University of Technology
Guangzhou, China

Abstract—Denial-of-Service (DoS) attacks especially distributed DoS (DDoS) attacks have become significant and increasing threats to the Internet. Huge efforts from both academia and industry have been made on detection and defense of DDoS attacks. However, most detection and defense schemes do not directly aim at protecting the victim of attacks itself (e.g., servers) but attack sources or intermediate network units. Although locating and identifying attacking sources are critical to stop attacks and for legal procedure, rapid and efficient predicting DDoS attacks to happen in the server is more important to reduce damage caused by attacks and even prevent attacks from happening. However, this part has not been addressed sufficiently in the literature. In this paper, we first briefly review research efforts on DDoS attacks, and then discuss a method to define and quantify attacks to servers based on available service rates. This is because the server is often the direct victim of DDoS attacks and the one-point failure of the entire service system. No matter whether there are attacks undergoing, if a server is overloaded even by normal service requests, the effect imposed to a service system is equivalent to that of attacks. Then a prediction method for the available service rate of the protected server is proposed, which applies the Auto Regressive Integrated Auto Regressive (ARIMA) model. Finally, we investigate the proposed prediction method to predict DDoS attacks through simulation studies with NS2. The simulation results show that the prediction algorithm is effective to predict most attacks.

Index Terms—Denial-of-service attacks (DoS), Distributed DoS (DDoS), available service rate, prediction-based detection of DDoS and low-rate TCP attacks.

I. INTRODUCTION

Unlike computer virus, Denial-of-service (DoS) attacks [11] do not aim at wrecking or stealing users’ data and information but disrupting or interrupting the service provided to legitimate users. DoS attacks aim at crippling applications, servers and even whole networks by disrupting or interrupting legitimate users’ communication. This kind of attacks always infinitely consume resources of networks, servers and end hosts with malicious objectives of stopping or degrading services to legitimate users. Typical resources targeted by such attacks include network bandwidth, server capacity, CPU capacity and memory (buffers) and interruption process capacity. For example, TCP SYN attacks consume the limited buffer of the TCP process and DNS flood attacks [1] utilize the weakness of the DNS protocol to generate high volumes of traffic to a targeted DNS server.

The most popular form of DoS attacks launched from the Internet is Distributed DoS (DDoS) attacks [2], [3]. DDoS employs dozens, hundreds or even over hundred thousands of compromised computers to perform a coordinated and widely distributed DoS attacks to quickly cause the victim (e.g., servers) to be broken, shut down or even rebooted. With the rapid development of the Internet technology, new approaches and tools to launch DDoS attacks are also advanced quickly so that DDoS attacks become easier and more automatic. Nowadays, almost an ordinary computer user can launch DDoS attacks as long as (s)he has the knowledge of the Internet and a little network operation skill. The worse is that DDoS tools can be easily downloaded from the Internet so that DDoS attack implementation is very simple and attacks can be easily launched by clicking the mouse. Although the operation and implementation of DDoS attacks are simple and easy, the damage caused by attacks is often fatal and disastrous.

Current research on DDoS attacks are focused on attack prevention, detection and response [6]. So far, lots of effort has been made on detection and defense of DDoS attacks. However, most schemes do not directly aim at protecting the victim (e.g., servers) but attack sources or intermediate network units [4]. Prevention of DDoS attacks often adopt techniques that preserve the integrity of hosts and detect rate-limit abnormal network activity [5]. Attack detection is typically based on methods such as signature matching [7] and anomaly detection [8]. Response to DDoS attacks typically involves filtering of attack packets by assuming a pre-defined signature of attacks and trace-back attempting to identify the attack paths [10]. To handle DDoS attacks efficiently, precise detecting attacks is the most essential and important phase, in which properly distinguishing legitimate traffic from malicious traffic is the most challenge job. Many DDoS attacks detecting methods have been proposed and most of them just detect certain pre-known attacks.

Although locating and identifying of attacking sources are critical to stop attacks and legal procedure, a rapid and efficient prediction of DDoS attacks to happen in the server is more important to reduce damage caused by attacks and even prevent attacks from happening. However, this part has not been
addressed sufficiently in the literature. In this paper, we discuss an approach to predict the service rate to be available in a server in order to prevent the server from being overloaded. With such a prediction, we can take some precaution measures in advance to prevent a crackdown of the server, which may be caused by DDoS attacks or other factors such as system malfunctions.

Obviously, predicting the premonition of DDoS attacks is important and effective to defend a server from DDoS attacks. The sampling metric to evaluate undergoing attacks is a critical factor. Common metrics used to evaluate server availability and service load include CPU occupancy and memory utilization as well as queue length. Here we induce available service rate as a measurement to qualify server’s availability in order to detect DDoS attacks. Then a prediction method based on the Auto Regressive Integrated Auto Regressive (ARIMA) model [13] for the available service rate is discussed by dividing server resources into CPU time, memory utilization and networking buffer. Based on the prediction, we use abnormal detecting technology to analyze the consumption of server resources to predict whether the server is under DDoS attacks. We also investigate the proposed prediction method through simulation studies with NS2. The simulation studies show that the prediction algorithm is effective to predict most attacks.

The remainder of this paper is organized as follows. In Section II, we introduce service rate to quantify server availability for attack detection. Then an available service rate prediction model is discussed based on the ARMA model in Section III. We verify the proposed prediction method through computer simulation with NS2 in Section IV. Finally, the paper is concluded in Section V.

II. PREDICTION MODEL FOR AVAILABLE SERVICE RATE

In this section, we first discuss the server model against attacks and then give a definition of available service rate to quantify server availability based on input loads. This quantity is used predict DDoS attacks discussed in the next section.

A. Server versus attacks

To quantify available service rate against process loads, defining and selecting proper audit metrics are critical for accurate prediction on server availability. For attack detection, there are many metrics such as queue throughput, the number of active TCP connections, distributions of source IP addresses, TCP buffer utilization and hop counts of packet’s journey. During a DDoS attack, thousands or even more packets moving through the network to the same server will shortly exhaust the limited resources such as radio links, processor memory and network bandwidth. On the other hand, A server can be abstracted by a queuing model with simple service discipline of first-come-first-service. Packets arrive at a server at random times. If a packet finds that the server is available, it will be immediately served for an amount of CPU processing time; otherwise, the packet will be enqueued. When a queue with a finite size becomes full, the newly arriving packets are dropped.

DoS attacks just aim at decreasing service efficiency or interrupting legitimate service by exhausting the limited resources. The metric selected here should reflect the availability of a server and explore the evidence of undergoing or potential attacks to detect DDoS attacks. As defined below, the available service rate can reflect the availability and efficiency of a server and is used to detect possible DDoS attacks according to the historical and current time series of available service rates.

Here, we model DDoS attacks of low-rate denial of service [12] which, unlike high-rate attacks, are difficult for routers and counter-DoS mechanisms to detect simply based on the flaws of TCP congestion control algorithms. As shown in Fig. 1, the queue consists of two input flows: attack packet flow and background traffic flow. Background traffic is the legitimate traffic under the normal network status, and attack traffic is added to model DDoS attacks. The queue input and service processes are stochastic.

![Model of DDoS attack to server](image)

**Fig. 1.** Model of DDoS attack to server

B. Available service rates

The instantaneous service rate of a server is approximated by the service rate averaged over a short time interval. Usually, a available service rate can be calculated by \( \rho = \frac{\mu}{T^\prime} \), where \( \lambda \) donates the sum of mean arrival rates of background and attack traffic, and \( \mu \) donates the mean service rate of the server. In order to consider the influence of past streams for prediction, we amend the above calculation as discussed below. Let \( T_0 \) denote the sample time interval. For the \( k \)-th time interval from \((k-1)T_0\) to \( kT_0 \), we sample the \( \lambda_k, \mu_k \) and \( q_{k-1} \), then available service rate during the \( k \)-th time interval is:

\[
\rho_k = \frac{\mu_k}{\lambda_k + q_{k-1}},
\]

where \( q_{k-1} = \lambda_{k-1} - \mu_{k-1} \) \((q_0 = 0)\) is the number of packets remaining in the queue after time \((k-1)T_0\). In this way, we can construct a series of available service rates against time \( \{ \rho_k \} \) as illustrated in Fig. 2.

III. AN ARIMA BASED PREDICTION ALGORITHM

From Section II-B, we can know a series of available service rates which can be built time by time. We assume that observations on available service rates take place at discrete time points with a fixed interval \( \Delta t \) (i.e., sample interval).
Apparently, the smaller \( \Delta t \), the closer the observation is to the actual one and the more accurate the prediction will be. Given \( N \) consecutive samples of available service rate available for analysis, namely, \( \rho_1, \rho_2, \ldots, \rho_t, \ldots, \rho_N \), which are observed at equidistant time points \( \tau_0 + \Delta t, \tau_0 + 2 \Delta t, \ldots, \tau_0 + i \Delta t, \ldots, \tau_0 + N \Delta t \). We can adopt \( \tau_0 \) as the start time and \( \Delta t \) as the time interval. Let \( Z \) denote a time series of available service rate, i.e., \( Z = \{ \rho_1, \rho_2, \ldots, \rho_t, \ldots, \rho_N \} \), and \( H_t = \{ h_1, h_2, \ldots, h_N \} \) an \( N \) samples observed before time \( t \), which is called historical series henceforth.

### A. Principle of prediction

Now we analyze the time series of available service rates for prediction. Theoretically given a time series \( \{ r_k \} \), when the number of samples is large enough and if the autocorrelation of \( \{ r_k \} \) does not follow rapid decreasing trend and never decreasing to zero as \( k \) increases, we can assert that this time series is non-stationary; otherwise, it is stationary. Given a discrete time series of samples \( H_t \), we can calculate the autocorrelation of this time series by

\[
r_k = \frac{1}{N} \sum_{i=0}^{N-k} (h_i - \bar{h})(h_{i+k} - \bar{h}), \quad k = 1, 2, \ldots, N, \tag{2}
\]

where \( \bar{h} \) is the mean of sample time series, and \( r_k \) is the autocorrelation of the sample value at \( k \). With some calculation, we can prove that this series is non-stationary. Therefore, traditional autoregressive (AR) and moving average (MA) processes are not suitable to analyze this series. Here we adopt the autoregressive integrated moving average model (ARIMA) [13] to model this series, which can consider the ‘non-stationary’ of this series properly.

The ARIMA model has three parameters and is expressed in ARIMA\((p, q, d)\), where \( p \) and \( q \) are the difference ranks respectively for autoregressive and average processes, and \( d \) is the difference rank to convert the non-stationary processes to stationary processes. From the DDoS attack model discussed in Section II-A, the service rate at time \( t \) is determined by the packets remaining in the queue at the last time so that both \( p \) and \( q \) are set to 1. When DDoS attacks arrive, the service rate will decrease linearly or exponentially. Here, the linear decreasing is taken so that \( d = 1 \). Therefore, ARIMA\((1,1,1)\) is adopted in this paper, whose differential equation is

\[
\nabla \rho_t - \phi \nabla \rho_{t-1} = a_t - (\theta)a_{t-1},
\]

\[
\nabla \rho_t = \rho_t - \rho_{t-1}, \tag{3}
\]

where \( a_t \) (\( \text{E}[a_t] = 0 \), \( \text{var}[a_t] = \sigma_a^2 \)) is a white noise process with mean zero and constant variance at \( t \). \( \phi \) and \( \theta \) are two constant parameters depending on the sample series. With a historical series before \( t \) and ARIMA\((1,1,1)\), by taking conditional expectations, we can predict the available service rate at time \( t \) through re-writing the model equation as follows:

\[
\hat{\rho}_t(1) = (1 + \phi)\rho_t - \phi \rho_{t-1} - \theta a_t, \quad \tag{4}
\]

\[
\hat{\rho}_t(l) = (1 + \phi)\hat{\rho}(l-1) - \phi \hat{\rho}(l-2), \quad \tag{5}
\]

where \( \hat{\rho}_t(l) \) is an \( l \) step prediction at time \( t + l \). So, if \( \phi, \theta \) and \( \sigma_a^2 \) are known, based on the available service rate \( \rho_t \) at the current time \( t \) and the last available service rate \( \rho_{t-1} \) at time \((t-1)\), we can calculate the future available service rate \( \rho_{t+1} \) at \((t+1)\). In this way, we can recursively predict a number of available service rates after time \( t \).

Since \( \phi, \theta \) and \( \sigma_a^2 \) are unknown, an initial estimation of these three parameters is discussed below. First, we calculate \( w_t = \nabla \rho_t = \rho_t - \rho_{t-1} \), \((t = 1, 2, \ldots, N)\) to form a time series \( \{ w_t \} \). We obtain the first three autocovariances \( c_0, c_1, c_2 \) of \( \{ w_t \} \) as follows:

\[
c_k = \frac{1}{N} \sum_{i=1}^{N-k} (w_i - \bar{w})(w_{i+k} - \bar{w}), \quad k = 0, 1, 2 \tag{6}
\]

According to the theory of the ARIMA prediction model, the estimation of \( \phi \) is \( \hat{\phi} = \frac{\bar{c}_1}{\bar{c}_0} \). Then, we calculate \( w'_t = w_t - \hat{\phi}w_{t-1} \), \((t = 1, 2, \ldots, N)\) to form \( \{ w'_t \} \). We obtain the first three autocovariances of \( \{ w'_t \} \) denoted as \( \bar{c}'_0, \bar{c}'_1, \bar{c}'_2 \). Then \( \bar{\theta} \) and \( \bar{\sigma}_a \), the estimations of \( \theta \) and \( \sigma_a \), can be calculated by

\[
\bar{\sigma}_a^2 = \frac{c'_0}{1 + \bar{\theta}^2}, \tag{7}
\]

\[
\bar{\theta} = -\frac{\bar{c}'_1}{\bar{\sigma}_a^2}. \tag{8}
\]

### B. Prediction-based detection algorithm

The flow chart of a prediction-based detection algorithm is shown in Fig. 3. The abnormality detection algorithm is based on series prediction on available service rates. The major steps of the proposed prediction-based detection algorithm are described below.

(a) Before the initial time \( t \), \( N \) \((N \geq 1)\) samples of available service rates are calculated according to (1) discussed in Section II to form a following historical time series before \( t \): \( H_t = \{ \rho_{t-i}, i \geq 1 \} = \{ h_1, h_2, \ldots, h_N \} \). Then

![Model for available service rates of the server](image)

![Detecting algorithm based on prediction on available service rate](image)

we can easily get the mean of this historical time series as 

$$E_k = E[H_t] = \frac{1}{N} \sum_{i=0}^{N} h_i. \quad (7)$$

(b) Based on $H_t$, according to (7) and (8) discussed in Section III-A, calculate the estimation of $\phi$, $\theta$ and $\sigma^2$, i.e., $\hat{\phi}$, $\hat{\theta}$ and $\hat{\sigma}^2$. We can obtain one step prediction following (4) and $l$ steps prediction with (5).

Now we get $K$ predictions at future time points at $t + 1, t + 2, \cdots, t + K$, which are denoted by $\hat{P}_t = \{\hat{\rho}_1, \hat{\rho}_2, \cdots, \hat{\rho}_K\}$. We can also get the mean of predicted time series $E_p = E[\hat{P}_t] = \frac{1}{N} \sum_{i=0}^{K} \hat{\rho}_i$ and the variation $\text{var}[\hat{P}_t] = \frac{1}{N} \sum_{i=0}^{K} (\hat{\rho}_i - E_p)^2$. If $E[\hat{P}_t] < \alpha E[H_t] \quad (0 < \alpha \leq 1)$, we define that there will be attacks in next interval time $T$, where $\alpha$ is an experiential constant meaning that the available service rate of a server will be evidently decreased to a fraction of the normal historical ones ($\alpha$). If $E[\hat{P}_t] \geq \alpha E[H_t]$, we cannot determine whether there is an attack and further learning is required by going to step (c).

(c) Calculate $K$ samples of available service rate for next time from $t + 1$ to $t + K$, which are denoted by $\hat{P}_t = \{\hat{\rho}_1, \hat{\rho}_2, \cdots, \hat{\rho}_K\}$, and the prediction average square error ($\xi_0$) as follows:

$$\xi_0 = \frac{1}{K} \sum_{i=1}^{K} (\hat{\rho}_i - \rho_i)^2. \quad (9)$$

If $||E_p - E_h|| > \delta$, where $\delta$ is a small positive value, it means that the prediction time series is not accurate enough. In this case, $H_t$ needs to be renewed by taking into account $\{\rho_1, \rho_2, \cdots, \rho_K\}$ so that $H_t = \{h_{N-K-1}, h_{N-K-2}, \cdots, h_N, \rho_1, \rho_2, \cdots, \rho_K\}$.

(d) Follow the algorithm in step (b) to predict $K$ available service rates from time $t + K + 1$ to $t + 2K$, which are denoted by $\hat{P}_{t+K} = \{\hat{\rho}_{K+1}, \hat{\rho}_{K+2}, \cdots, \hat{\rho}_{2K}\}$. Then the available service rate at $t + K$, $P_{t+K} = \{\rho_{K+1}, \rho_{K+2}, \cdots, \rho_{2K}\}$ and the second round prediction average square error ($\xi_1$) is

$$\xi_1 = \frac{1}{K} \sum_{i=1}^{K} (\hat{\rho}_{K+i} - \rho_{K+i})^2. \quad (10)$$

If $\xi_1 > \beta \xi_0$ ($\beta > 1$), it means that there will be a great change in the next time interval and DDoS streams are included in the traffic; otherwise, continue detecting and sampling. After certain times, we get the total prediction average square error with (10).

IV. SIMULATION STUDIES

In this section, we investigate the performance of the prediction-based detection algorithm discussed in Section II through computer simulation IN NS-2 [14]. We simulate low-rate TCP attacks [12] and generate attack $S$ in the network as illustrated in Fig. 4, where node $N$ is the upstream router. The link between $N$ and $S$ is a bottleneck link with a link rate equal to 2Mbit/s. $\{N_1, \cdots, N_5\}$ are the normal nodes in the network, which generate normal traffic (i.e., background traffic) exponentially with a mean rate equal to 800kB/s and both idle and burst times equal to 500ms. Nodes $A_1, \cdots, A_m$ are abnormal nodes compromised by attackers, whose traffic forms DDoS streams with a high rate of 10Mbit/s and idle time equal to 500ms while burst time equal to 100ms.

We simulated distributed low-rate DoS attacks three times and each lasts 10 seconds. In the first 10 seconds, there are just 5 normal nodes $\{N_1, \cdots, N_5\}$ sending normal traffic to server $S$. At a random time $10i + t_1$ ($i = 1, 2, 3$), where $t_1$ ($0 < t_1 < 1$) is a random variable. We select randomly 10 nodes from $\{A_1, \cdots, A_m\}$ to send hundreds of TCP low-rate flows to server $S$. Eventually, at $10 + t_i$, the packet loss rate is increased rapidly and the available service rate is reduced quickly.

Using the algorithm discussed in Section II, with $\Delta t = 0.05$ second and $N = 60$ with sampling 5000 available service rates, we can predict the available service rate of the protected server in the future. As shown in Fig. 5 where the solid line indicates the actual available service rate while the dashed line with cross is the predicted value of the available service rate. Overall, only 16 predictions are much different from the actual results, resulting a prediction average square error of $9.76 \times 10^{-3}$. Furthermore, we can find that the predicted lines near DDoS attack time are different from the actual available service rates since the prediction is based on the historical records without attacks. However, when an attack is coming, there may be a change point near these special times, which are also useful for attack detection.

Table I shows the prediction accuracy against and the number of historical samples ($N$) and sample intervals ($\Delta t$) adopted in the prediction. We can find that the larger $N$ or the shorter $\Delta t$, the more accurate the predicted results will be. We can also find that the relative errors of for most predictions are around $1 \sim 3\%$.

V. CONCLUSION

In this paper we discussed a prediction-based detection algorithm against Distributed Denial-of-Service (DDoS). The proposed prediction algorithm adopts historical available service rates of a server to predict the server availability in the
future by using the autoregressive integrated moving average model (ARIMA). Based on this prediction, we can detect abnormal states of the protected server, which might be caused by ongoing DDoS attacks. In fact, this prediction algorithm tries to alarm any possible abnormality of a server in terms of its available service rate in the future. Our simulation results indicate that this prediction algorithm can provide predicted results well matching the simulation results.

However, using prediction for DDoS attacks has been seldom mentioned in the literature, and there are many issues to be addressed further such as the determination of empirical parameter in the model proposed in this paper. In the future, we will try to expand this approach to detect other kinds of attacks while further improvement on the current algorithm will also be carried out.

REFERENCES


Fig. 5. Prediction of available service rates versus the actual results

### TABLE I

<table>
<thead>
<tr>
<th>N (Δt = 0.05 second)</th>
<th>10</th>
<th>20</th>
<th>40</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error (10⁻³)</td>
<td>6.25</td>
<td>5.8</td>
<td>4.21</td>
<td>3.95</td>
</tr>
<tr>
<td>Relative error (%)</td>
<td>1.46</td>
<td>1.35</td>
<td>0.98</td>
<td>0.91</td>
</tr>
<tr>
<td>Δt (second, N = 60)</td>
<td>10</td>
<td>20</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>Error (10⁻³)</td>
<td>2.57</td>
<td>3.59</td>
<td>3.59</td>
<td>11.9</td>
</tr>
<tr>
<td>Relative error (%)</td>
<td>0.61</td>
<td>0.83</td>
<td>0.91</td>
<td>2.77</td>
</tr>
</tbody>
</table>

---


