Available online at www.sciencedirect.com
ScienceDirect
Computer Networks 51 (2007) 3564–3573
Computer Networks
www.elsevier.com/locate/comnet

# A queueing analysis for the denial of service (DoS) attacks in computer networks

Yang Wang [a], Chuang Lin [a], Quan-Lin Li [b], Yuguang Fang [c,*]

[a] Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China
[b] Department of Industrial Engineering, Tsinghua University, Beijing 100084, China
[c] Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611, United States

Received 10 May 2006; received in revised form 25 December 2006; accepted 19 February 2007
Available online 18 March 2007

Responsible Editor: Christos Douligeris

**Abstract**

In most network security analysis, researchers mainly focus on qualitative studies on security schemes and possible attacks, and there are few papers on quantitative analysis in the current literature. In this paper, we propose one queueing model for the evaluation of the denial of service (DoS) attacks in computer networks. The network under DoS attacks is characterized by a two-dimensional embedded Markov chain model. With this model, we can develop a memory-efficient algorithm for finding the stationary probability distribution which can be used to find other interesting performance metrics such as the connection loss probability and buffer occupancy percentages of half-open connections for regular traffic and attack traffic. Different from previous works in the literature, this paper gives a more general analytical approach to the study of security measures of a computer network under DoS attacks. We hope that our approach opens a new avenue to the quantitative evaluation of more complicated security schemes in computer networks.


*Keywords:* DoS attack; Network security; Queueing; Connection loss probability

## 1. Introduction

Due to the widely deployed wide-area computer and communications networks, the Internet has undergone rapid development all over the world and has become indispensable in our daily lives. However, the Internet has caused many security problems and financial loss due to the unauthorized access. Network security has thus attracted considerable attention in the last few decades. The first international well-publicized security incident of the ARPANET was identified by Cliff Stoll in 1986 [1]. Later in 1988, the ARPANET had experienced the first automated network security attack, referred to as the Morris worm [1]. As network capability grows faster and larger, network security has become a rather important issue from both theoretical point of view and engineering applications.

* Corresponding author. Tel.: +1 352 846 3043.
*E-mail addresses:* ywang@csnet1.cs.tsinghua.edu.cn (Y. Wang), clin@csnet1.cs.tsinghua.edu.cn (C. Lin), liquanlin@tsinghua.edu.cn (Q.-L. Li), fang@ece.ufl.edu (Y. Fang).

1389-1286/$ - see front matter 
doi:10.1016/j.comnet.2007.02.011

Network attacks are common nowadays. There are several types of crucial attacks, such as the denial of service (DoS), worm, trojan horse and virus, each of which causes serious problems to normal business operations. The DoS attacks usually cause significant disruptions to computer networks. A DoS attack can be regarded as an explicit attempt of attackers to prevent legitimate users from gaining a normal network service (see Sandstrom [2]). Due to the serious consequence of DoS attacks, there are intensive research in this area. In general, DoS attacks can be classified into several different types, however, the prevalent type is referred to as the packet flooding attack. Attackers may flood a network with a large volume of data in order to deliberately consume the basic and limited resources of a victim, such as the process control blocks and the maximum allowed connections. In particular, DoS attacks may disrupt the normal operation of physical components in the network, and may also manipulate data in transit such as encrypted data [3]. Moreover, multiple hosts may be employed to coordinate an attack by flooding a victim with a barrage of attack packets, which is usually referred to as the distributed denial of service (DDoS) attacks. A reflection attack, which is a special case of DDoS attacks, uses the compromised hosts as reflectors to hide the identity of the attackers or to amplify an attack. Therefore, the reflection attacks can cause more severe damages to the networks.

Many defense mechanisms have been proposed in the literature to defend against DoS attacks [4–8]. For these mechanisms, most researchers primarily focused on attack detections and responses. Commonly, the anomaly-detection or signature-scan technique can be used to identify an ongoing attack; while the responses can take the reactive or proactive measures to mitigate the networking damages. These mechanisms include blocking attack packets to reduce the intensity of attacks, tracing the packets to locate the attacking source(s), and using the proactive measures to filter the attack packets ([4] and reference therein). Based on the effective detection and filtering techniques, we may be able to characterize the behavior of the DoS attacks and estimate the impact of the DoS attacks quantitatively. Moore et al. [9] provided some insights on the prevalence of DoS activities over the Internet and used a traffic monitoring technique, called *backscatter analysis*, to estimate the worldwide prevalence of the DoS attacks. Hussain et al. [6] proposed a framework for classifying the DoS attacks based on the header contents, the transient ramp-up behaviors and the spectral analysis, and presented a statistical analysis for the DoS attacks and came up with some defense mechanisms, a useful study of quantifying the dynamics of DoS attacks.

However, only a few available works have employed rigorous mathematical models to analytically study the DoS attacks. This motivates us in this paper to develop some more general mathematical models (such as queueing models) to analyze the DoS attacks. Along this line, Chang [4] mentioned a simple queueing model for the SYN flooding attack, which is one of the most common DoS attacks. Long et al. [7] proposed two queueing models for the DoS attacks in order to obtain the packet delay jitter and the loss probability. Khan and Traore [8] analyzed the impact of DoS attacks on three parameters: the arrival rate, the queue-growth-rate, and the response time, which were used for the attack detection. Huang et al. [10] used a generalized multi-class Erlang and Engset mixed loss model to analyze the DDoS attacks, which was later extended for 3G wireless cellular networks [11]. Different from these works in the literature, our paper studies the DoS attacks analytically by using a more general queue model, a two-dimensional embedded Markov chain, which can more accurately capture the dynamics of the actual DoS attacks.

The main contributions of this paper are twofold. The first one is the theoretical model for the study of the DoS attacks, which is motivated by a few available results on quantitative analysis of the DoS attacks. We propose a two-dimensional queueing model to evaluate the system performance of a computer network under DoS attacks. The queueing model of this paper is novel because we model the system with two requesting queues for regular requests and the attack requests with different service time distributions. In this model, all connection requests share the same backlog queue, and each request immediately receives a buffer space of the backlog queue once it arrives at the system upon finding an idle buffer space and is blocked otherwise. The difficulty in the analysis lies in the fact that the regular requests and the attack requests behave differently, e.g., their buffer occupancy times will have different probability distributions. Therefore, we can only use the two-dimensional embedded Markov chain to characterize the DoS attacks. The second contribution of the paper is to compute the stationary probability distribution of the embedded Markov chain, from which we can com-

pute some key security metrics, such as the connection loss probability. Based on the theoretical results on the security metrics, we can set up some crucial parameters such as the buffer size and the holding time for half-open connections in order to guarantee the desired degree of service availability under certain attack scenarios.

The organization of this paper is as follows. In Section 2, we present our queueing model for the system under the SYN flooding attacks. In Section 3, we give the two-dimensional embedded Markov chain for the queueing model for the analysis of the DoS attacks. A memory effective level-eliminating algorithm for the computation of the stationary probability distribution is developed in Section 4. In Section 5, we discuss a few security metrics of interest and use numerical examples to illustrate the system performance. Finally, we conclude our paper in Section 6.

## 2. Model description

To intuitively describe our modeling approach, in this section we construct a queueing model for the SYN flooding attack and characterize it using a two-dimensional embedded Markov chain. In our model, we focus on the consumption of the limited resources, which is a common and crucial characteristic for various DoS attacks. It is worthwhile to note that this approach can be extended to model other DoS attacks, though the details of various models may be quite different.

It is well known that SYN flooding attacks exploit network vulnerabilities with respect to the TCP protocol, where the three-way handshake algorithm is used. In general, the arrival of SYN packets contains two types: the regular request packets and the attack packets that request for connections. A large number of SYN packets are always sent to a victim for pretending to make connections with the victim. However, the victim can hardly differentiate the attack packets from the regular request packets, and therefore it has to respond by sending back the SYN-ACK packets. The nodes with regular request packets will respond in a timely fashion while the attacking nodes sending out attack packets will not respond, leaving half-open connections in the victim's backlog queue for a period. For the attack packets, the victim may keep one half-open connection in the buffer of the backlog queue for the incomplete handshake until a response is dropped, although the victim may possibly retransmit the SYN–ACK packets for several times. It is clear that these half-open connections can quickly consume all the memory allocated for the pending connections and prevent the victim from further accepting new requests, leading to the well-known buffer overflow problem. Due to this consideration, some operating systems have to control the buffer space and/or the holding time with respect to the half-open connections. The purpose is to defend against the SYN flooding attacks as effectively as possible.

For the SYN flooding attacks, we now construct a queueing model to analyze the buffer occupancy. We assume that each half-open connection is held for at most a deterministic period of time $b$, which is a time interval from the epoch that the half-open connection begins to the epoch that the connection is dropped. The victim has a connection buffer of the backlog queue, in which at most $N$ half-open connections are allowed simultaneously. Specifically, we assume that the half-open connection for a regular request packet is held for a random time which is exponentially distributed with parameter $\mu$. The arrivals of the regular request packets and the attack packets are both Poisson processes with rates $\lambda_1$ and $\lambda_2$, respectively. The two arrival processes are independent of each other and of the holding times for half-open connections.

Based on the above assumptions, the system under SYN flooding attacks can be modeled as a two-dimensional queueing model with $N$ servers, two arrival processes and two service times of different distribution. In what follows, we will analyze this queueing system.

## 3. Two-dimensional embedded Markov chain

In this section, we derive the analytical results for the two-dimensional embedded Markov chain for our problem.

Let $N_1(t)$ and $N_2(t)$ be the numbers of the regular request packets and the attack packets at time $t$, respectively. It is obvious that $N_1(t) + N_2(t) \leqslant N$ for all $t \geqslant 0$. Let $\tau_n$ be the $n$th time that the half-open connections of the $n$th attack packet is dropped. It is clear that $\{N_1(\tau_n), N_2(\tau_n){:}n \geqslant 0\}$ is a two-dimensional Markov chain with the state space given by

$$\Omega = \{(n_1, n_2) : n_1 \geqslant 0, n_2 \geqslant 0, n_1 + n_2 \leqslant N\}.$$

For simplicity, we divide the state space $\Omega$ into $N+1$ levels, where level $i = \{(i,j), 0 \leqslant j \leqslant N - i\}$ for $0 \leqslant i \leqslant N$, where $i$ and $j$ denote $i$ regular request packets and $j$ attack packets in the buffer, respectively.

To find the transition probability matrix of the Markov chain, we first introduce some notations as follows. For $1 \leqslant m \leqslant N$ and $k \geqslant 0$,

$$F_m(x) = 1 - \mathrm{e}^{-m\mu x},$$
$$a_k = \int_0^{+\infty} \mathrm{e}^{-\lambda_1 x} \frac{(\lambda_1 x)^k}{k!}\, \mathrm{d}F_1(x),$$
$$\bar{a}_k = \sum_{l=k}^{+\infty} a_l,$$
$$a_k^{(m)} = \int_0^{b} \mathrm{e}^{-\lambda_1 x} \frac{(\lambda_1 x)^k}{k!}\, \mathrm{d}F_m(x),$$
$$\bar{a}_k^{(m)} = \sum_{l=k}^{+\infty} a_l^{(m)},$$
$$a^{*(m)} = \sum_{k=0}^{+\infty} \int_b^{+\infty} \mathrm{e}^{-\lambda_1 x} \frac{(\lambda_1 x)^k}{k!}\, \mathrm{d}F_m(x),$$
$$b_k = \mathrm{e}^{-\lambda_2 b} \frac{(\lambda_2 b)^k}{k!},$$
$$\bar{b}_k = \sum_{l=k}^{\infty} b_l.$$

For the aforementioned queueing model, there are two correlated queueing lines dealing with the regular request packets and the attack packets, respectively. The transition probability matrices corresponding to the following two queues can be easily written according to Chapter 1 of Neuts [12]. The details are omitted below. In what follows we describe each of them using the levelling techniques.

1. *The queue for the regular request packets*
When there are $i$ attack packets in this system for $0 \leqslant i \leqslant N$, the regular request packets are described as an $M_1/M/N-i/N-i$ queue. In this case, the Markov chain related to the service time $b$ of the $M_1/M/N-i/N-i$ queue has the transition probability matrix of size $(N-i+1)\times(N-i+1)$ as follows:

$$Q_1^{(i)} = \begin{pmatrix} a_0 & a_1 & a_2 & a_3 & \cdots & a_{N-i-1} & \bar{a}_{N-i} \\ a_0^{(1)} + a^{*(1)} & a_1^{(1)} & a_2^{(1)} & a_3^{(1)} & \cdots & a_{N-i-1}^{(1)} & \bar{a}_{N-i}^{(1)} \\ a^{*(2)} & a_0^{(2)} & a_1^{(2)} & a_2^{(2)} & \cdots & a_{N-i-2}^{(2)} & \bar{a}_{N-i-1}^{(2)} \\ a^{*(3)} & & a_0^{(3)} & a_1^{(3)} & \cdots & a_{N-i-3}^{(3)} & \bar{a}_{N-i-2}^{(3)} \\ \vdots & & & \ddots & & \vdots & \vdots \\ a^{*(N-i-1)} & & & & & a_1^{(N-i-1)} & \bar{a}_2^{(N-i-1)} \\ a^{*(N-i)} & & & & & a_0^{(N-i)} & \bar{a}_1^{(N-i)} \end{pmatrix}. \tag{1}$$

2. *The queue for the attack packets*
When there are $j$ regular request packets in this system for $0 \leqslant j \leqslant N$, the attack packets can be described as an $M_2/D/N-j/N-j$ queue. In this case, the Markov chain of the $M_2/D/N-j/N-j$ queue, according to (5.5.53) of Neuts [12], has the probability transition matrix of size $(N-j+1)\times(N-j+1)$ denoted as below:

$$Q_2^{(j)} = \begin{pmatrix} b_0 & b_1 & & b_{N-j-1} & \bar{b}_{N-j} \\ b_0 & b_1 & \cdots & b_{N-j-1} & \bar{b}_{N-j} \\ \vdots & \vdots & & \vdots & \vdots \\ b_0 & b_1 & & b_{N-j-1} & \bar{b}_{N-j} \end{pmatrix}. \tag{2}$$

3. *The composite and dependent queue*
Based on the two queueing models given in (1) and (2), we can model the DoS attacks as a two-dimensional embedded Markov chain with a probability transition matrix given by

$$P = \begin{pmatrix} P_{0,0} & P_{0,1} & P_{0,2} & P_{0,3} & \cdots & P_{0,N-1} & P_{0,N} \\ P_{1,0} & P_{1,1} & P_{1,2} & P_{1,3} & \cdots & P_{1,N-1} & P_{1,N} \\ P_{2,0} & P_{2,1} & P_{2,2} & P_{2,3} & \cdots & P_{2,N-1} & P_{2,N} \\ \vdots & & P_{3,2} & P_{3,3} & & \vdots & \vdots \\ P_{N-1,0} & & & \ddots & & P_{N-1,N-1} & P_{N-1,N} \\ P_{N,0} & & & & & P_{N,N-1} & P_{N,N} \end{pmatrix}, \tag{3}$$

where $P_{i,i}$ is the probability transition matrix, each entry of which is the probability transition between two states in level $i$; while $P_{i,j}$ is a probability transition matrix, each entry of which is the probability transition from a state of level $i$ to a state of level $j$. The size of $P_{i,i}$ is $(N+1-i)\times(N+1-i)$ and the sizes of the other submatrices can be determined accordingly.

The entries of submatrices $P_{i,j}$ can be accurately computed from $Q_1^{(i)}$ and $Q_2^{(j)}$ according to the following three types.

*Type one: The diagonal submatrices*

$$P_{0,0} = a_0 \begin{pmatrix} b_0 & b_1 & & b_{N-1} & \bar{b}_N \\ b_0 & b_1 & \cdots & b_{N-1} & \bar{b}_N \\ \vdots & \vdots & & \vdots & \vdots \\ b_0 & b_1 & & b_{N-1} & \bar{b}_N \end{pmatrix}, \tag{4}$$

$$P_{i,i} = a_1^{(i)} \begin{pmatrix} b_0 & b_1 & & b_{N-i-1} & \bar{b}_{N-i} \\ b_0 & b_1 & \cdots & b_{N-i-1} & \bar{b}_{N-i} \\ \vdots & \vdots & & \vdots & \vdots \\ b_0 & b_1 & & b_{N-i-1} & \bar{b}_{N-i} \end{pmatrix},$$

$$1 \leqslant i \leqslant N-1, \tag{5}$$

$$P_{N,N} = \bar{a}_1^{(N)}. \tag{6}$$

*Type two: The lower-triangle submatrices*

$$P_{1,0} = \begin{pmatrix} a_0^{(1)} + a^{*(1)} & & & & & \\ & a_0^{(1)} + a^{*(1)} & & & & \\ & & a_0^{(1)} + a^{*(1)} & & & \\ & & & \ddots & & \\ & & & & a_0^{(1)} + a^{*(1)} & 0 \end{pmatrix},$$

$$P_{i,i-1} = \begin{pmatrix} a_0^{(i)} & & & & \\ & a_0^{(i)} & & & \\ & & a_0^{(i)} & & \\ & & & \ddots & \\ & & & a_0^{(i)} & 0 \end{pmatrix}, \quad 2 \leqslant i \leqslant N,$$

$$P_{i,0} = \begin{pmatrix} a^{*(i)} & & & & & \\ & a^{*(i)} & & & & \\ & & a^{*(i)} & & & \\ & & & \ddots & & \\ & & & a^{*(i)} & 0 & \cdots & 0 \end{pmatrix}, \quad 2 \leqslant i \leqslant N.$$

*Type three: The upper-triangle submatrices*

$$P_{0,j} = \begin{pmatrix} a_j & & & & \\ & a_j & & & \\ & & a_j & & \\ & & & \ddots & \\ & & & & \bar{a}_j \\ & & & & 0 \\ & & & & 0 \\ & & & & \vdots \\ & & & & 0 \end{pmatrix}, \quad 1 \leqslant j \leqslant N-1$$

$$P_{i,j} = \begin{pmatrix} a_{j-i+1}^{(i)} & & & & & & \\ & a_{j-i+1}^{(i)} & & & & & \\ & & a_{j-i+1}^{(i)} & & & & \\ & & & \ddots & & & \\ & & & & \bar{a}_{j-i+1}^{(i)} & & \\ & & & & 0 & & \\ & & & & \vdots & & \\ & & & & 0 & & \end{pmatrix},$$

$$1 \leqslant i \leqslant N-2, \quad i+1 \leqslant j \leqslant N-1,$$

$$P_{0,N} = (\bar{a}_N \quad 0 \quad 0 \quad \cdots \quad 0)^{\mathrm{T}},$$

$$P_{i,N} = \left(\bar{a}_{N-i+1}^{(i)} \quad 0 \quad 0 \quad \cdots \quad 0\right)^{\mathrm{T}}, \quad 1 \leqslant i \leqslant N-1.$$

We observe that the blocks of the probability matrix $P$ given in (3) have a similar form to that in (1), while each diagonal submatrix in (4)–(6) has a similar form to that in (2). This is why we write down (1) and (2) to help the understanding of the complicated structure of the matrix $P$.

## 4. A level-eliminating algorithm

In this section, we propose a memory effective algorithm to compute the stationary probability distribution of the two-dimensional Markov chain, which will be very useful for calculating the security measures.

Since the matrix $P$ only has the finite dimension and is irreducible, it is positive recurrent. Let $\pi = (\pi_0, \pi_1, \ldots, \pi_{N-1}, \pi_N)$ be the stationary probability distribution of $P$, partitioned according to the $N+1$ different levels. Further, we write $\pi_i = (\pi_{i,0}, \pi_{i,1}, \ldots, \pi_{i,N-i})$ for $0 \leqslant i \leqslant N$.

Since $\pi$ is the stationary probability distribution of $P$, we have

$$\pi P = \pi \tag{7}$$

and

$$\pi e = 1, \tag{8}$$

where $e$ is a column vector of ones.

It follows from (7) that

$$\pi(I - P) = 0. \tag{9}$$

Notice that the system of linear Eqs. (8) and (9) has finite dimension, thus some algorithms given in the literature can be used to solve the system of linear equations, such as Gauss elimination. However, they have to deal with higher-dimensional matrices computations, especially when $N$ is large, hence in general are not efficient for the model of this paper, because it is not always easy and even difficult to deal with the computation of higher-dimensional matrices directly due to memory requirement. In order to efficiently solve the system of Eqs. (8) and (9), we propose a novel algorithm: a level-eliminating method in this section. Based on the specific structure of the matrix $P$ (i.e., the structure of the systems of the equations we have), this new algorithm can avoid the calculation of high dimensional matrices by decomposing them into small ones. This new algorithm is based on the censoring technique [14–17] and is to eliminate $N$ these levels of matrix $P$ from level 0, to level 1, to level $2, \ldots,$ up to level $N$, step by step, until the final matrix is censored to only one level.

To facilitate this process, we introduce a sequence of matrices $\{P^{(i)}, 0 \leqslant i \leqslant N\}$, where $P^{(0)} = I - P$ and $P^{(i)}$ is the probability transition matrix of the $i$th censoring Markov chain in which level 0, level 1, ..., level $i-1$ have already been censored to levels $1-N$, levels $2-N$, ..., levels $i-N$, respectively.

The level-eliminating method is given as follows:

*Step one*: Initialization.

Let $P^{(0)} = I - P$ and $\pi^{(0)} = \pi$.

*Step two*: Censoring process.

Censor the matrix $P^{(0)}$ to the matrix $P^{(1)}$ where level *0* is eliminated, and continue until we obtain the matrix sequence $\{P^{(i)}\}$. The stationary probability distribution $\pi^{(i)}$ satisfies

$$\pi^{(i)} P^{(i)} = 0. \tag{10}$$

We now present the concrete eliminating process. We partition the probability matrix $P^{(i)}$ into four submatrices as follows:

$$P^{(i)} = \begin{pmatrix} T^{(i)} & U^{(i)} \\ V^{(i)} & Q^{(i)} \end{pmatrix},$$

where $T^{(i)}$ denotes the upper-left block submatrix of the $P^{(i)}$ with the same dimension of probability transition matrix at level $i$.

We also partition the stationary probability distribution $\pi^{(i)}$ into two sub-vectors $\pi^{(i)} = (\pi_1^{(i)}, \pi_2^{(i)})$, according to the partition of $P^{(i)}$. Then the Eq. (10) can be rewritten as

$$\left(\pi_1^{(i)}, \pi_2^{(i)}\right) \begin{pmatrix} T^{(i)} & U^{(i)} \\ V^{(i)} & Q^{(i)} \end{pmatrix} = 0.$$

Solving the matrix equation, we obtain

$$\pi_1^{(i)} T^{(i)} + \pi_2^{(i)} V^{(i)} = 0 \tag{11}$$

and

$$\pi_1^{(i)} U^{(i)} + \pi_2^{(i)} Q^{(i)} = 0. \tag{12}$$

Notice that the matrices $T^{(i)}$ and $Q^{(i)}$ are all invertible, thus it follows from (11) that

$$\pi_1^{(i)} = -\pi_2^{(i)} V^{(i)} \left(T^{(i)}\right)^{-1}. \tag{13}$$

Substituting (13) into (12), we obtain

$$\pi_2^{(i)} \left(Q^{(i)} - V^{(i)} (T^{(i)})^{-1} U^{(i)}\right) = 0. \tag{14}$$

Thus, if we let

$$P^{(i+1)} = Q^{(i)} - V^{(i)} (T^{(i)})^{-1} U^{(i)} \tag{15}$$

and

$$\pi^{(i+1)} = \pi_2^{(i)}, \tag{16}$$

then, we have

$$\pi^{(i+1)} P^{(i+1)} = 0. \tag{17}$$

*Step three*: Calculating the stationary probability vectors.

Using the Eq. (13), we can iteratively represent the vector sequence $\{\pi_1^{(i)},\ 0 \leqslant i \leqslant N\}$. Following is the detailed algorithm:

For $i = N - 1$ down to *0*,
  $\pi_1^{(i)} = -\pi^{(i+1)} V^{(i)} (T^{(i)})^{-1}$;
  $\pi_i = \pi_1^{(i)}$;
  $\pi^{(i)} = (\pi_1^{(i)}, \pi^{(i+1)})$;
End

Since $\pi = (\pi_0, \pi_1, \ldots, \pi_{N-1}, \pi_N)$, $\pi^{(i)} = (\pi_1^{(i)}, \pi_2^{(i)})$, $\pi^{(i+1)} = \pi_2^{(i)}$, and the size of $\pi_1^{(i)}$ is the same size of probability transition matrix at level $i$, it is easy to see that $\pi_i = \kappa \pi_1^{(i)}$, where the normalization condition $\pi e = 1$ leads to $\kappa = 1/\Sigma_{i=0}^{N} \pi_1^{(i)} e$.

Now, we compare the complexity of our level-eliminating algorithm with the known algorithms, such as the Gauss elimination algorithm, which usually has to carry out calculations on higher-dimensional matrices. Since it is well-known that the calculation of higher-dimensional matrices is rather complex and sometimes even harder to implement due to the computational limitations and memory requirements, our new algorithm can decompose a higher-dimensional matrix into some lower-dimensional matrices, and perform calculation on small matrices instead. For the DoS model, the matrices involving in our algorithm have the dimension $N+1$ at most while the number is much smaller than $N(N+1)/2$, the dimension of the matrices for the Gauss elimination algorithm. Thus our algorithm is more memory efficient and can be extended to modeling of larger size. On the other hand, the time complexity of our level-eliminating algorithm is mostly based on the iterative computations of (15). As commonly used in the literature, we only need to estimate the number of multiplications involved. From Eq. (15), the number of multiplications for our algorithm is $O(\sum_{i=1}^{N+1} i^5/4) = O(N^6)$, while that for the Gauss elimination algorithm is also $O(N^6)$. We see that our algorithm has at most the same complexity as the Gaussian elimination algorithm. Furthermore, since most of the calculations in our algorithm are matrix multiplications, we can use the fast parallel matrix multiplication

strategy [13], which requires $O(m^{2.376})$ operations for an arbitrary matrix of dimension $m$. In this way, the time complexity of our algorithm can be reduced to $O(N^{5.7})$.

## 5. Security performance metrics and numerical examples

In this section, we derive some important security metrics which characterize the performance of the network system under DoS attacks. First, the connection loss probability is an important measure in the connection depletion DoS attacks. Moreover, people also care for the buffer occupancy percentages of half-open connections for regular traffic and attack traffic, which can be given as the average ratios of the numbers of the regular half-connections $N_1$ and the attack packets $N_2$ to the maximum allowable number of half-open connections $N$, respectively. These two metrics represent how severe the DoS attacks deteriorate the system performance. In this section, we also give some numerical examples to illustrate how to quantify these security metrics. Numerical examples indicate that the analytical method of this paper is effective and efficient in the study of network security.

Connection loss probability is a basic measure for assessing the performance of the network under DoS attacks. Since each arriving packet must be dropped once there have already been $N$ pending connections in the system, the connection loss states can be specified as the pairs $(N_1, N_2)$ with $N_1 + N_2 = N$. If there is no more room for pending connections to be saved, the arriving packets have to be blocked. Using the stationary probability distribution given in Section 4, the connection loss probability can be described as

$$P_{\text{loss}} = \sum_{i=0}^{N} \pi_{i,N-i}. \tag{18}$$

Obviously, if the connection loss probability $P_{\text{loss}}$ is large, the network should be under DoS attack. From this understanding, we may use a threshold value $\delta > 0$ small enough to indicate the network security status: if

$$P_{\text{loss}} < \delta,$$

we can conclude that the network is not under DoS attack.

The way that DoS attacks work is to send as many attack packets as possible to consume network resources (i.e., buffer spaces) such that $P_{\text{loss}} \geqslant \delta$, which leads to network performance degradation.

Fig. 1 shows the connection loss probability $P_{\text{loss}}$ of the system with respect to the attack traffic load with different settings for the maximum allowable number of half-open connections $N$ and holding time $b$ for pending connections. Let $\lambda_1 = 10/\text{s}$ as the parameter for the Poisson arrival process of the regular request packets. Let $\lambda_2 = k\lambda_1$, and clearly, the parameter $k$ may be understood as the ratio of arrival rates between the attack packets and the regular request packets. For simplicity, we use the exponential distribution with the parameter $\mu = 100/\text{s}$ as the service time of regular request packets, and it could represent the severity of congestions in the network. We change $k$ from 0.1 to 2, which indicate the attack traffic load from low to high. Let $N = 20$ and $b = 5$ as the basic parameters. From Fig. 1 we easily observe that the packet loss probability $P_{\text{loss}}$ increases with the increase of the attack traffic load. Either decreasing the holding time $b$ or increasing the maximum allowable number of half-open connections $N$ can reduce $P_{\text{loss}}$ if

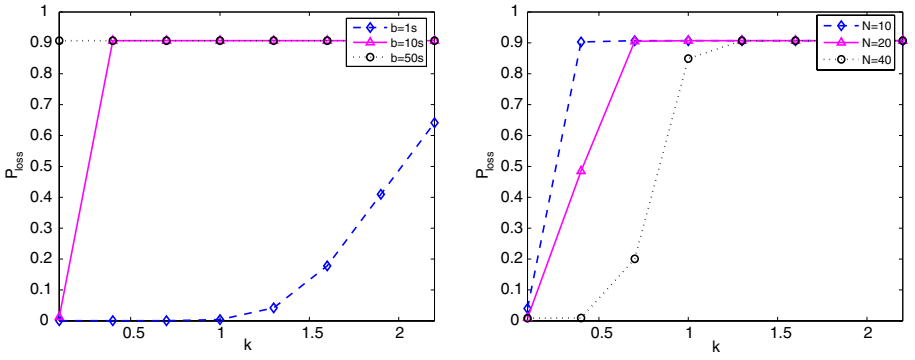



Fig. 1. The connection loss probability vs. the attack load for different parameters.

the attack traffic load is limited. Nevertheless, when the traffic is already under heavy load, there are not much of difference.

Next, we consider the buffer occupancy percentage of half-open connections for regular traffic, which is characterized by the mean ratio of the number of regular half-open connections to the maximum allowable number of half-open connections, to be denoted as $P_r$. We can obtain

$$P_r = \frac{1}{N}\sum_{i=0}^{N} i \sum_{j=0}^{N-i} \pi_{i,j}. \tag{19}$$

Similarly, the buffer occupancy percentage of half-open connections for attack packets, which is represented by the mean ratio of the number of attack packets to the maximum allowable number of half-open connections, denoted by $P_a$, can also be obtained as:

$$P_a = \frac{1}{N}\sum_{i=0}^{N} \sum_{j=0}^{N-i} j\pi_{i,j}. \tag{20}$$

Figs. 2 and 3, respectively, depict how the buffer occupancy percentages of half-open connections for regular traffic and attack traffic depend on the attack traffic load with different system parameter $b$ and $N$. We also assume $\lambda = 10/\text{s}$ and $\mu = 100/\text{s}$ as the general parameters. We can easily observe that $P_r$ almost remains at a similar level in the two figures of Fig. 2, which implies that it is not sensitive to $b$ and $N$. The buffer occupancy percentage of half-open connections for attack packets is much larger than the regular request packets, because the holding time of attack packets is much longer. It is clear that $k$ is a crucial factor, and the attack load will severely degrade the performance of the network under DoS attack.

Note that the numerical examples are to validate the DoS model in Section 2 and the algorithm given in Section 4, and also show the analysis of security metrics. So here we choose the proper $N$ in our numerical examples for sake of simplicity. Our level-eliminating algorithm can quickly solve the systems of linear equations of the DoS model for

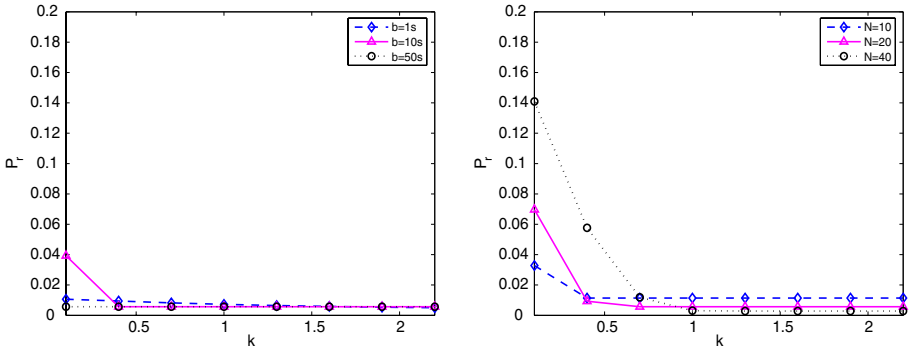



Fig. 2. The mean ratio of the number of regular half-connections to the maximum allowable number of half-open connections vs. the attack load for different parameters.

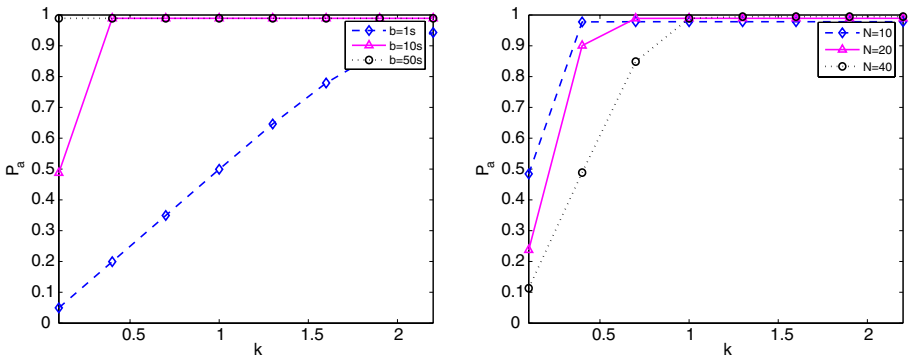



Fig. 3. The mean ratio of the number of attack packets to the maximum allowable number of half-open connections vs. the attack load for different parameters.

$N \leqslant 100$, as indicated in our examples. As $N$ gets larger, we can use parallel techniques to implement our algorithm efficiently.

Since the time complexity of our level-eliminating algorithm is mostly based on the iterative computation of (15), we can use the typical parallel matrix multiplication algorithms to efficiently solve the larger models. Introducing the parallel technique, we can use multiple computing nodes to solve the model, e.g. using a 32-node cluster or grid systems. There the matrix can be partitioned into blocks which can be computed separately on each node.

The theoretical analysis of this paper proposes a new avenue to the quantitative evaluation of computer networks under DoS attacks. Following this way, simulations of the realistic DoS attacks are crucial work to facilitate applying the analysis results to our real world. Simulation by our work is not only to validate the model, but the much more important is to apply the results in designing efficient attack detect and defense mechanisms. From the simulation in practical terms, we first can get the main parameters to characterize the systems status under DoS attacks, which are to describe attributes of security. Adjusting the system parameters according to the theoretical analysis, we could then find a proper setting for the security mechanism of the system and further design new protection strategies. In addition, some improvements of the model and analysis method can be developed with respect to the practical systems. Obviously, it needs much work to do and we would like to practice that in the future work.

## 6. Conclusions and future work

In this paper, we analytically characterize the security attributes of a network under the DoS attacks such as the most prevalent SYN-flooding attack. We use a two-dimensional embedded Markov chain to find the stationary probability distribution and find analytical results for a few security performance metrics. This paper represents a more comprehensive work in the quantitative study of network security under the DoS attacks and may open a new avenue to assess the security issue of more complex networks under the DoS attacks.

We hope that the research insights gained by this paper would lead to practical deployment and applications to some crucial domains of network security. There are still some issues to be dealt with in the future. First, it is necessary to provide some effective simulations of the real DoS attacks. In the literature, the general datasets for the DoS attacks do not provide the fine details to find the probability distributions. We should generate and analyze the datasets to construct such probability distributions. Based on this, we could find the characteristics of the DoS and DoS-like attacks, estimate the effect of the attacks quantitatively, design new detections and protections in computer networks. Second, it is a key point to improve the model description and algorithms. We will further strengthen numerical simulations and practical deployments for the study of network security.

## Acknowledgements

The work of Wang and Lin was supported in part by the National Natural Science Foundation of China under Grant Nos. 90412012 and 60673187 and the National Grand Fundamental Research 973 Program of China under Grant No. 2006CB708301. The work of Li was supported in part by the National Natural Science Foundation of China under Grant No. 10671107 and the National Grand Fundamental Research 973 Program of China under Grant No. 2006CB805901. The work of Fang was supported in part by the US National Science Foundation under grant CNS-0626881 and under CAREER Award ANI-0093241.

## References

[1] T. Longstaff, J. Ellis, S. Hernan, H. Lipson, R. McMillan, L. Pesante, D. Simmel, Security of the Internet, The Froehlich/Kent Encyclopedia of Telecommunications, vol. 15, Marcel Dekker, New York, 1997, pp. 231–255.

[2] H. Sandstrom, A survey of the denial of service problem, BSc Programmes in Engineering Computer Engineering, 2001.

[3] L. Meyer, W.T. Penzhorn, Denial of service and distributed denial of service – today and tomorrow, in: IEEE AFRICON, 2004, pp. 959–964.

[4] R.K.C. Chang, Defending against flooding-based distributed denial-of-service attacks: a tutorial, IEEE Communications Magazine 40 (10) (2002) 42–51.

[5] H. Wang, D. Zhang, K. Shin, Detecting SYN flooding attacks, in: Proceedings of IEEE INFOCOM 2002, pp. 1530–1539.

[6] A. Hussain, J. Heidemann, C. Papadopoulos, A framework for classifying denial of service attacks, in: Proceedings of SIGCOMM, 2003, pp. 99–110.

[7] M. Long, C.H. Wu, J.Y. Hung, Denial of service attacks on network-based control systems: impact and mitigation, IEEE Transactions on Industrial Informatics 1 (2) (2005) 85–96, May.

[8] S. Khan, I. Traore, Queue-based analysis of DoS attacks, in: Proceeding of the 2005 IEEE Workshop on Information

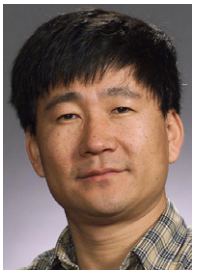
Assurance and Security, United States Mulitary Academy, West Point, NY, pp. 266–273.
[9] D. Moore, G. Voelker, S. Savage, Inferring internet denial of service activity, in: Proceeding of the USENIX Secority Symposium, Washington, DC, USA, August 2001, pp. 9–22.
[10] Q. Huang, H. Kobayshi, B. Liu, Analysis of a new form of distributed denial of service attack, in: Proceedings of the Conference on Information Science and Systems, Johns Hopkins University, 2003, pp. 12–14, March.
[11] Q. Huang, H. Kobayshi, B. Liu, Modeling of distributed denial of service attacks in wireless networks, in: IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, vol. 1, 2003, pp. 41–44.
[12] M.F. Neuts, Structured Stochastic Matrices of M/G/1 Type and Their Applications, Marcel Dekker Inc., New-York and Basel, 1989.
[13] D. Coppersmith, S. Winograd, Matrix multiplication via arithmetic progressions, Journal of Symbolic Computation 9 (1990) 251–280.
[14] Q.L. Li, J.H. Cao, Two types of RG-factorizations of continuous-time level dependent quasi-birth-and-death processes and their applications to stochastic integral functionals, Stochastic Models 20 (3) (2004) 299–340.
[15] Q.L. Li, Y.Q. Zhao, The RG-factorization in block-structured Markov renewal processes with applications, in: Xun Zhu (Ed.), Observation, Theory and Modeling of Atmospheric Variability, World Scientific, 2004, pp. 545–568.
[16] Q.L. Li, Y.Q. Zhao, Light-tailed asymptotics of stationary probability vectors of Markov chains of GI/G/1 type, Advances in Applied Probability 37 (4) (2005) 075–1093.
[17] Q.L. Li, Y.Q. Zhao, Heavy-tailed asymptotics of stationary probability vectors of Markov chains of GI/G/1 type, Advances in Applied Probability 37 (2) (2005) 482–509.

**Yang Wang** received the B.E. degree in Computer Science and Technology from Tsinghua University, China in 2004. Since 2004, she has been a Ph.D. student at Tsinghua University. Her research interests are in the areas of network security and wireless networks.

**Chuang Lin** is a Professor and the Head of the Department of Computer Science and Technology, Tsinghua University, Beijing, China. He received a Ph.D. degree in Computer Science from Tsinghua University in 1994. His current research interests include computer networks, performance evaluation, network security analysis, and Petri net theory and its applications. He has published three books and more than 250 papers in research journals and IEEE conference proceedings.

He is a senior member of the IEEE and the Chinese Delegate in TC6 of IFIP. He served as the General Chair, ACM SIGCOMM Asia Workshop 2005. He is currently the Associate Editor for IEEE Transactions on Vehicular Technology, the Area Editor for Journal of Parallel and Distributed Computing, and the Area Editor for Journal of Computer Networks.

**Quan-Lin Li** is an Associate Professor at the Department of Industrial Engineering, Tsinghua University, Beijing, PR China. He received a Ph.D. degree from the Institute of Applied Mathematics, Chinese Academy of Sciences, Beijing, PR China in 1998. His research interests include stochastic models, stochastic processes, stochastic process algebra, manufacturing systems, communication networks, and network security. He has published over 30 papers in international research journals such as, Advances in Applied Probability, Queueing Systems and Stochastic Models.



**Yuguang Fang** received a Ph.D. degree in Systems Engineering from Case Western Reserve University in January 1994 and a Ph.D. degree in Electrical Engineering from Boston University in May 1997. He was an Assistant Professor in the Department of Electrical and Computer Engineering at New Jersey Institute of Technology from July 1998 to May 2000. He then joined the Department of Electrical and Computer Engineering at University of Florida in May 2000 as an assistant professor, got an early promotion to an associate professor with tenure in August 2003 and to a full professor in August 2005. He holds a University of Florida Research Foundation (UFRF) Professorship from 2006 to 2009. He has published over 200 papers in refereed professional journals and conferences. He received the National Science Foundation Faculty Early Career Award in 2001 and the Office of Naval Research Young Investigator Award in 2002. He is the recipient of the Best Paper Award in IEEE International Conference on Network Protocols (ICNP) in 2006 and the recipient of the IEEE TCGN Best Paper Award in the IEEE High-Speed Networks Symposium, IEEE Globecom in 2002. He has served on several editorial boards of technical journals including IEEE Transactions on Communications, IEEE Transactions on Wireless Communications, IEEE Transactions on Mobile Computing and ACM Wireless Networks. He has also been actively participating in professional conference organizations such as serving as The Steering Committee Co-Chair for QShine, the Technical Program Vice-Chair for IEEE INFOCOM'2005, Technical Program Symposium Co-Chair for IEEE Globecom'2004, and a member of Technical Program Committee for IEEE INFOCOM (1998, 2000, 2003–2007).