Hindawi Publishing Corporation
Journal of Sensors
Volume 2015, Article ID 465402, 11 pages
http://dx.doi.org/10.1155/2015/465402

Research Article

Accurately Identifying New QoS Violation Driven by High-Distributed Low-Rate Denial of Service Attacks Based on Multiple Observed Features

Jian Kang,[1,2] Mei Yang,[3] and Junyao Zhang[4]

[1] *Department of Computer Science & Technology, Jilin University, Changchun 130012, China*
[2] *Key Laboratory of Symbol Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun 130012, China*
[3] *Department of Software Engineering, Jilin University, Changchun 130012, China*
[4] *Department of EECS, University of Central Florida, Orlando, FL 32816, USA*

Correspondence should be addressed to Jian Kang; kj885788@gmail.com

Received 4 August 2014; Revised 24 November 2014; Accepted 8 December 2014

Academic Editor: Jun Zhang

We propose using multiple observed features of network traffic to identify new high-distributed low-rate quality of services (QoS) violation so that detection accuracy may be further improved. For the multiple observed features, we choose *F feature* in TCP packet header as a microscopic feature and, *P feature* and *D feature* of network traffic as macroscopic features. Based on these features, we establish *multistream fused hidden Markov model* (MF-HMM) to detect stealthy low-rate denial of service (LDoS) attacks hidden in legitimate network background traffic. In addition, the threshold value is dynamically adjusted by using Kaufman algorithm. Our experiments show that the additive effect of combining multiple features effectively reduces the false-positive rate. The average detection rate of MF-HMM results in a significant 23.39% and 44.64% improvement over typical power spectrum density (PSD) algorithm and nonparametric cumulative sum (CUSUM) algorithm.

1. Introduction

In recent years malicious quality of services (QoS) violation attacks have become one of the most serious security threats to the Internet. New QoS attacks are increasingly showing the trend of high-distributed low rate. In the literature, this kind of attacks has been called *shrew attacks* [1], *pulsing denial of service (DoS) attacks* [2], or *reduction of quality (RoQ) attacks* [3]. For simplicity, we call all of them *LDoS (low-rate denial of service) attacks* in the sequel.

LDoS attacks are stealthy, periodic, pulsing, and low rate in attack volume, very different from early flooding type of attacks. A traditional detection system against flooding attacks is based on traffic volume analysis method in the time domain. However, it almost has no effect on new LDoS attack [4]. This is because the average bandwidth consumption differs very little between normal and attack streams.

In this paper, we present a new approach to identify LDoS attacks by combining multiple observed features at the micro- and macrolevel. Multidimensional features are extremely valuable for describing slight changes of network properties and help us accurately differentiate attack flows. So our new approach can complement existing detection mechanisms based on one-dimensional feature and overcome the bottleneck of detection accuracy for LDoS violation.

In microscopic features, we calculate *weighted summation of flag bits* (WSFB) in TCP packet header to reflect the packet's internal slight change with and without LDoS attacks. Macroscopically, the best distinguishing characteristic between LDoS and normal flow is different periodicity in frequency domain [5]. Based on this fact, we choose *weighted average size of packet in queue* (WASPQ) in router as an observed sequence. Then, we convert the WASPQ sequence into frequency-domain spectrum using discrete

Fourier transform (DFT) and achieve the power spectrum density (PSD) of WASPQ as a macroscopic feature. Moreover, we calculate the *difference between request/response flows* (DRRF) as another macroscopic feature.

Based on above three-dimensional features, we develop a multistream fused hidden Markov model (MF-HMM) to detect LDoS violation hidden in legitimate TCP/IP traffic. In addition, we adjust the *decision threshold* value dynamically based on Kaufman algorithm for improving the detection accuracy. Notations, symbols, and abbreviations used in this paper are summarized in Notations section. Only brief definitions are given here; details are given in subsequent sections.

The rest of this paper is organized as follows. In Section 2, we present the related work. Section 3 describes MF-HMM, its advantages, and its training algorithm. Section 4 presents the overview of TF-HMM procedure and explains how to extract multiobserved features of network traffic to establish the corresponding component HMM of TF-HMM. Furthermore, we also introduce the threshold dynamic adjustment based on Kaufman algorithm. In Section 5, we compare our work with those of other researchers and discuss the training and recognition time of TF-HMM. Finally, we conclude our paper in Section 6.

2. Related Work

Some scholars studied the mathematical model of LDoS attacks. By simulating various LDoS attacks, they discussed the properties of LDoS attacks and gave some suggestions on further research. Maciá-Fernández et al. [6] summarized the behavior of LDoS and proposed a mathematical model for the LDoS attack. They also discussed the development trend and made some recommendations for building defense techniques against this attack. He et al. [7] presented theoretical analyses, modeling, and simulations of various LDoS attacks. And they discussed the difficulties of defending and current solutions. Zhu et al. [8] discussed the vulnerabilities of TCP and the principle of low-rate attacks. Moreover, the simulation of attacks was investigated, and the further direction of research is suggested.

Most current LDoS-related studies focus on using the frequency domain method to detect LDoS attack and have made clear progress. A research group [9] proposed an approach of detecting LDoS attack based on the model of small signal. Furthermore, in paper [10], they presented the method of multiple sampling averaging based on missing sampling (MSABMS) to detect LDoS attacks. An eigenvalue-estimating matrix was established to estimate the attack period after the detection of LDoS attacks. In addition, they also indicated a scheme [11] of detecting LDoS attack based on time window sampling in time domain and capturing the periodicity by statistic analysis in frequency domain. Zhang et al. [12] proposed a detection method, which is similar to that of Yu et al. [13]. In this method, the sum of the power spectrum is computed within 1–50 Hz, and the intersection of the two fitting curves is taken as the judging threshold. Luo and Chang [2] proposed a two-stage scheme to detect LDoS attacks on a victim network. The first stage is a discrete wavelet transform (DWT) analysis of the network traffic. The second stage is to detect change points by using a non-parametric cumulative sum (CUSUM) algorithm. Liu [14] proposed an LDoS attack detection method by calculating the *Holder* based on binary discrete wavelet analysis. Shevtekar et al. [15] presented an approach of detecting the periodicity of attack flow based on autocorrelation of flow.

Some detection methods based on traditional traffic characteristics are proposed in recent years. These methods detect the LDoS attacks by searching and identifying the abnormal network traffic caused by the LDoS attacks. For example, the exponentially weighted moving average (EWMA) method was presented in papers [16, 17]. However, the EWMA algorithm may smooth not only the normal traffic but also the abnormal traffic. This will affect the detection accuracy for the LDoS attacks. Therefore, paper [18] proposed an adaptive EWMA method which used an adaptive weighting function instead of the constant weighting of EWMA algorithm. The adaptive EWMA can smooth the accidental error and retain the exceptional mutation. Thus, it is more efficient than EWMA method.

Unlike a popular deployment location of detection system, paper [19] proposed an adaptive detection method for LDoS attacks in *source-end* network. The method does not require the distribution assumption of the traffic samples. Moreover, they presented the automatic adjustment of the detection threshold according to the traffic conditions.

In particular, Xiang et al. [20] innovatively propose using two new information metrics to detect low-rate DDoS attacks by measuring the difference between legitimate traffic and attack traffic. The proposed *generalized entropy metric* and *information distance metric* outperform the existing popular approach as they can clearly enlarge the adjudication distance and then obtain the better detection sensitivity.

In summary, most researches use one-dimensional information of network traffic to establish algorithms for detecting LDoS attack. Though some algorithms are sophisticated, one-dimensional information is not enough to accurately differentiate stealthy LDoS attack hidden in legitimate traffic. Despite gratifying progress, the high false-positive rate is still a striking bottleneck.

3. Multistream Fused HMM

We first describe basic properties of multistream fused HMM and then give its mathematical description and training algorithm in detail.

3.1. Basic Properties. To accurately identify stealthy LDoS violation hidden in legitimate network traffic, the combination of multiobserved features is considered in our scheme by using multistream fused HMM [21]. According to the maximum entropy principle and the maximum mutual information (MMI) criterion, MF-HMM constructs a new structure linking multiple HMMs. MF-HMM is the generalization of two-stream fused HMM [22].

The main advantages of MF-HMM are as follows.

(1) Every observation feature can be modeled by a component HMM, so the performance of every feature can be analyzed individually. The set of features can be modified according to the performance analysis.

(2) Compared with other existing model fusion methods (e.g., CHMM [23], MHMM [24], etc.), MF-HMM reaches a better balance between model complexity and performance.

(3) MF-HMM has stronger robustness. If one component HMM fails due to some reason, the other component HMM can still work. Thus, the final result is still a valuable reference for the recognition judgment.

3.2. Mathematical Description. HMM is the basis of MF-HMM. In brief, we only discuss MF-HMM, and paper [25] discussed the HMM definition and relevant algorithms in detail. The mathematical symbols in this paper are consistent with the standard HMM description symbol.

Let $\{O^{(i)}, i = 1, \ldots, n\}$ represent n tightly coupled observing sequences. Assume that $\{O^{(i)}, i = 1, \ldots, n\}$ can be modeled by n corresponding HMMs with hidden states $\{Q^{(i)}, i = 1, \ldots, n\}$. In MF-HMM, an optimal solution for $p(O^{(1)}; O^{(2)}; \ldots; O^{(n)})$ is given according to the maximum entropy principle and the maximum mutual information criterion $\widehat{p}(O^{(1)}; O^{(2)}; \ldots; O^{(n)})$.

In order to calculate $\widehat{p}(O^{(1)}; O^{(2)}; \ldots; O^{(n)})$, firstly we need to calculate every component $\widehat{p}^{(i)}(O^{(1)}; O^{(2)}; \ldots; O^{(n)})$; here $i = 1, 2, \ldots, n$. The ith $\widehat{p}^{(i)}(O^{(1)}; O^{(2)}; \ldots; O^{(n)})$ can be given through

$$\begin{aligned}
&\widehat{p}^{(i)}\left(O^{(1)}; O^{(2)}; \ldots; O^{(n)}\right) \\
&\quad = p\left(O^{(1)}\right) p\left(O^{(2)}\right) \cdots p\left(O^{(n)}\right) \\
&\qquad \cdot \frac{p\left(Q^{(i)}, O^{(1)}, \ldots, O^{(i-1)}, O^{(i+1)}, \ldots, O^{(n)}\right)}{p\left(Q^{(i)}\right) p\left(O^{(1)}\right) \ldots p\left(O^{(i-1)}\right) p\left(O^{(i+1)}\right) \ldots p\left(O^{(n)}\right)} \\
&\quad = p\left(O^{(i)}\right) p\left(O^{(1)}, \ldots, O^{(i-1)}, O^{(i+1)}, \ldots, O^{(n)} \mid Q^{(i)}\right).
\end{aligned} \tag{1}$$

And assume

$$\begin{aligned}
&p\left(O^{(1)}, \ldots, O^{(i-1)}, O^{(i+1)}, \ldots, O^{(n)} \mid Q^{(i)}\right) \\
&\quad = \prod_{j \neq i, j=1}^{n} p\left(O^{(j)} \mid Q^{(i)}\right).
\end{aligned} \tag{2}$$

It has a good record in recognizing and detecting LDoS attacks, though the conditional independence assumption is always violated in practice. The success is because of the small number of parameters to be estimated in assumption. Without this assumption, some complicated algorithms require more training data and are more susceptible to local maximum during parameter estimation.

So, the estimate of $\widehat{p}^{(i)}(O^{(1)}; O^{(2)}; \ldots; O^{(n)})$ can be given by

$$\widehat{p}^{(i)}\left(O^{(1)}; O^{(2)}; \ldots; O^{(n)}\right) = p\left(O^{(i)}\right) \prod_{j \neq i, j=1}^{n} p\left(O^{(j)} \mid Q^{(i)}\right). \tag{3}$$

There are different expressions to different i. To our three-stream fused HMM (TF-HMM), (3) corresponds to (4a), (4b), and (4c) as follows;

$$\begin{aligned}
&\widehat{p}^{(1)}\left(O^{(1)}; O^{(2)}; O^{(3)}\right) \\
&\quad = p\left(O^{(1)}\right) p\left(O^{(2)} \mid Q^{(1)}\right) p\left(O^{(3)} \mid Q^{(1)}\right)
\end{aligned} \tag{4a}$$

$$\begin{aligned}
&\widehat{p}^{(2)}\left(O^{(1)}; O^{(2)}; O^{(3)}\right) \\
&\quad = p\left(O^{(2)}\right) p\left(O^{(1)} \mid Q^{(2)}\right) p\left(O^{(3)} \mid Q^{(2)}\right)
\end{aligned} \tag{4b}$$

$$\begin{aligned}
&\widehat{p}^{(3)}\left(O^{(1)}; O^{(2)}; O^{(3)}\right) \\
&\quad = p\left(O^{(3)}\right) p\left(O^{(1)} \mid Q^{(3)}\right) p\left(O^{(2)} \mid Q^{(3)}\right).
\end{aligned} \tag{4c}$$

In practice, if the n component HMMs have different reliabilities, they may be combined by different weights to get a better result:

$$\widehat{p}\left(O^{(1)}; O^{(2)}; \ldots; O^{(n)}\right) = \sum_{i=1}^{n} \lambda^{(i)} \widehat{p}^{(i)}\left(O^{(1)}; O^{(2)}; \ldots; O^{(n)}\right). \tag{5}$$

Here, $\sum_{i=1}^{n} \lambda^{(i)} = 1$.

3.3. Training Algorithm. The training algorithm of MF-HMM is a three-step process.

(1) n component HMMs are trained independently by representative algorithm, such as Baum-Welch algorithm, segmented K-means algorithm, or hybrid method EM algorithm.

(2) The best hidden state sequences of the component HMMs are estimated by the Viterbi algorithm.

(3) Calculate the coupling parameters between the n HMMs.

To our three-stream fused HMM, step (1) is to calculate (6a), (6b), and (6c):

$$\widehat{\Pi}^{(1)}, \widehat{A}^{(1)}, \widehat{B}^{(1)} = \arg \max_{\Pi^{(1)}, A^{(1)}, B^{(1)}} \left(\log p\left(O^{(1)}\right)\right) \tag{6a}$$

$$\widehat{\Pi}^{(2)}, \widehat{A}^{(2)}, \widehat{B}^{(2)} = \arg \max_{\Pi^{(2)}, A^{(2)}, B^{(2)}} \left(\log p\left(O^{(2)}\right)\right) \tag{6b}$$

$$\widehat{\Pi}^{(3)}, \widehat{A}^{(3)}, \widehat{B}^{(3)} = \arg \max_{\Pi^{(3)}, A^{(3)}, B^{(3)}} \left(\log p\left(O^{(3)}\right)\right). \tag{6c}$$

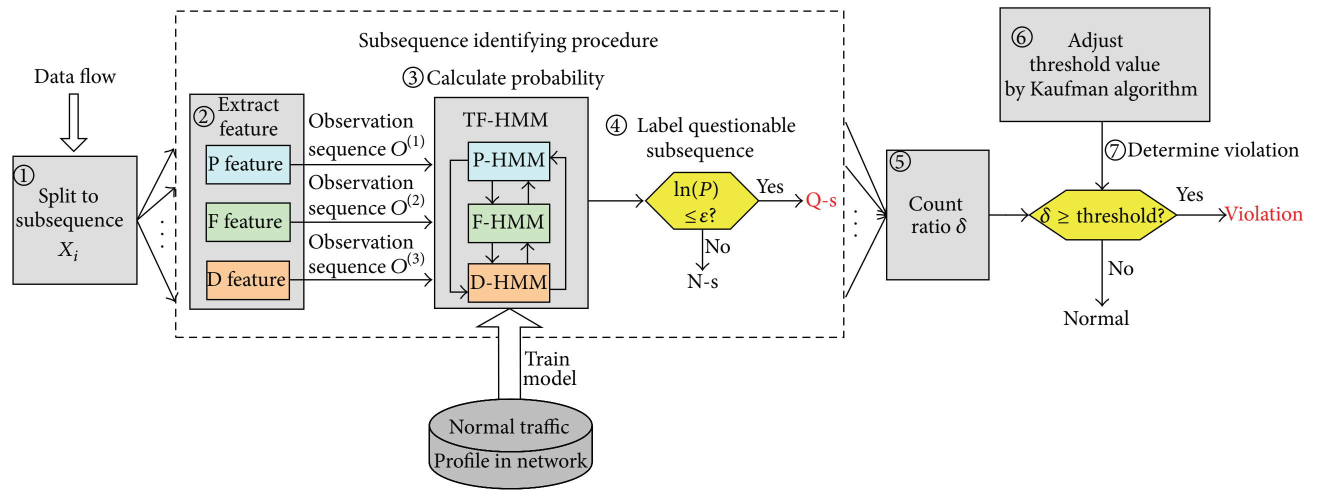

Data flow
① Split to subsequence X_i
Subsequence identifying procedure
③ Calculate probability
② Extract feature
P feature
F feature
D feature
Observation sequence $O^{(1)}$
Observation sequence $O^{(2)}$
Observation sequence $O^{(3)}$
TF-HMM
P-HMM
F-HMM
D-HMM
④ Label questionable subsequence
$\ln(P)$ $\leq \varepsilon$?
Yes
Q-s
No
N-s
Train model
Normal traffic
Profile in network
⑤ Count ratio δ
⑥ Adjust threshold value by Kaufman algorithm
⑦ Determine violation
$\delta \geq$ threshold?
Yes
Violation
No
Normal

Figure 1: Procedure of TF-HMM.

Step (2) is to calculate (7a), (7b), and (7c):

$$\widehat{Q}^{(1)} = \arg\max_{Q^{(1)}} \left(\log p\left(O^{(1)}, Q^{(1)}\right)\right) \tag{7a}$$

$$\widehat{Q}^{(2)} = \arg\max_{Q^{(2)}} \left(\log p\left(O^{(2)}, Q^{(2)}\right)\right) \tag{7b}$$

$$\widehat{Q}^{(3)} = \arg\max_{Q^{(3)}} \left(\log p\left(O^{(3)}, Q^{(3)}\right)\right). \tag{7c}$$

Step (3) is to estimate the coupling parameters between HMM1, HMM2, and HMM3:

$$\widehat{B}^{(1,2)} = \arg\max_{B^{(1,2)}} p\left(O^{(2)} \mid \widehat{Q}^{(1)}\right) \tag{8a}$$

$$\widehat{B}^{(1,3)} = \arg\max_{B^{(1,3)}} p\left(O^{(3)} \mid \widehat{Q}^{(1)}\right) \tag{8b}$$

$$\widehat{B}^{(2,1)} = \arg\max_{B^{(2,1)}} p\left(O^{(1)} \mid \widehat{Q}^{(2)}\right) \tag{8c}$$

$$\widehat{B}^{(2,3)} = \arg\max_{B^{(2,3)}} p\left(O^{(3)} \mid \widehat{Q}^{(2)}\right) \tag{8d}$$

$$\widehat{B}^{(3,1)} = \arg\max_{B^{(3,1)}} p\left(O^{(1)} \mid \widehat{Q}^{(3)}\right) \tag{8e}$$

$$\widehat{B}^{(3,2)} = \arg\max_{B^{(3,2)}} p\left(O^{(2)} \mid \widehat{Q}^{(3)}\right). \tag{8f}$$

4. Identifying LDoS Violation Using TH-HMM

In this section, we first present the procedure of identifying LDoS violation by using TF-HMM. Then, we explain how to establish three-component HMMs of TF-HMM, including F-HMM, P-HMM, and D-HMM. At last, we introduce the threshold dynamic adjustment based on Kaufman algorithm.

4.1. Procedure Overview. In order to make it easier to understand, we firstly introduce the procedure of TH-HMM, as illustrated in Figure 1.

(1) Split into Subsequence. Let the length of the detected sequence be L. Split the detected sequence with a k length splitting window, so the set of these subsequences is $\{X_i\}$; here, $1 \leq i \leq L/k$.

(2) Extract Three Observed Features. Extract F feature, P feature, and D feature, and then form the three-dimensional observation state sequence.

(3) Calculate the Output Probability. Input state sequences to TF-HMM, and calculate the output probability $\ln \widehat{p}(O^{(1)}; O^{(2)}; O^{(3)})$ of every subsequence, denoted by $\ln(P)$.

(4) Label a Questionable Subsequence. If $\ln(P)$ is less than the threshold ε, it is labeled as a questionable subsequence (Q-s); otherwise it is marked as a normal subsequence (N-s).

(5) Count the Ratio of Questionable Subsequence. After computing and labeling all subsequences, count the ratio δ according to

$$\delta = \frac{\text{the number of questionable subsequences}}{\text{the total of all subsequences}}. \tag{9}$$

(6) Adjust Threshold Value by Kaufman Algorithm. During the detection system run, the threshold value will be adjusted by using Kaufman algorithm. In practice, the average detection rate of TF-HMM has been effectively improved with it.

(7) Determine the Violation. At last, compare δ with the decision threshold value *threshold*: if $\delta > \textit{threshold}$, it is determined as LDoS violations; else, there is no violations.

4.2. Establishing Three-Component HMMs. In order to apply TF-HMM, we extract multiobserved features of network traffic, including WSFB feature, PSD of WASPQ feature, and DRRF feature. They constitute three-dimensional observation state sequence. Each sequence is modeled by a component HMM. Three-component HMMs together make up TF-HMM.

Table 1: Weight of different flag bits in TCP header.

U R G	A C K	P S H	R S T	S Y N	F I N
2^5	2^4	2^3	2^2	2^1	2^0

4.2.1. F-HMM. In order to reduce network QoS, spoofed TCP/IP packets must be used. In microscopic view, attackers usually use random number to fill internal attribute fields of the forged packet, resulting in vast differences with the real data packet. We choose flag bits in TCP packet header as a microscopic feature to describe a slight internal change of packet attribute fields.

To enlarge differences of flag bits between the forged packets with the real ones, we define different weights to different flag bits [26], as in Table 1.

Next, we achieve the *weighted summation of flag bits* (WSFB) by using

$$O_{\text{wsfb}} = 2^5 \times \text{URG} + 2^4 \times \text{ACK} + 2^3 \times \text{PSH} + 2^2 \times \text{RST} + 2^1 \times \text{SYN} + 2^0 \times \text{FIN}. \tag{10}$$

So we can construct a component HMM based on the observing sequence of WSFB; simply mark it as F-HMM.

4.2.2. P-HMM. Paper [27] indicates that attack data packet occupies a certain proportion in router buffer queue at LDoS attack, and the greater the damage is, the higher the proportion is. At the same time, paper [28, 29] concludes that attackers must use the data packet as short as possible to achieve a good attack effect, which results in an obvious decrease of the average size of packets in buffer queue under attacks than under normal conditions. We introduce the *weighted average size of packet in queue* (WASPQ) to describe this periodicity change in macroscopic view.

Let the number of packets in queue when at sampling time t be N_t and let each size of packet be S_i, $i \in [1, N]$. In order to highlight the characteristic that the shorter the packet, the more important, we introduce weight γ_i, $\gamma_i \in [0, 1]$, and calculate the WASPQ value S_{WASPQ} as follows:

$$S_{\text{WASPQ}} = \frac{\sum_{i=1}^{N_t} \gamma_i S_i}{N_t}. \tag{11}$$

In order to depict inherent periodic feature of LDoS attack, we take S_{WASPQ} as the discrete signal series and sample it with a period of 0.1 sec. The change of S_{WASPQ} value with and without attacks is modeled by a random process: $\{s(t), t = n\Delta, n \in N\}$, where Δ is a constant time interval, which we assume 0.1 sec, and N is a set of positive integers, and, at each time point t, $s(t)$ is a random variable, representing the total number of S_{WASPQ} in $(t - \Delta, t]$.

To study the periodicity embedded in the $s(t)$ sequence, we use its autocorrelation function in discrete time as follows:

$$R_{xx}(m) = \frac{1}{N - m} \sum_{n=0}^{N-m+1} [s(n)\, s(n+m)]. \tag{12}$$

The $R_{xx}(m)$ captures the correlation of the $s(t)$ sequence and itself at interval m. If there is any periodicity existing, autocorrelation function is capable of finding it.

To figure out the periodicity embedded in the $s(t)$ sequence, we convert the autocorrelation time series by discrete Fourier transform (DFT) to generate the power spectrum density (PSD) as follows:

$$\text{PSD}(f) = \text{DFT}(R_{xx}(m), f) = \frac{1}{N} |X[f]|^2, \tag{13}$$

where $X[f] = \sum_{n=0}^{N-1} R_{xx}(m) \exp(-j2\pi fn/N)$ is the N-point DFT, $f = 0, 1, 2, \ldots, N - 1$.

We note that we use the standard periodogram rather than Welch's method of averaged periodogram [30]. This is because in our work we are interested in the detection and estimation of a single periodic feature, which is better achieved using the standard periodogram as discussed in [31].

Therefore, we can get the component HMM based on the PSD of WASPQ feature, simply referred to as P-HMM.

4.2.3. D-HMM. In a normal TCP session of two-way communications, the request flow is limited by the response flow [32]. In the macroscopic view, the difference value between them should remain relatively stable normally. In case of LDoS attacks, a huge number of forgery request packets will lead to a sharp increase of the difference. Therefore, we introduce the *difference between request/response flows* (DRRF) to represent the difference change.

Let the sequence $d[i]$ be the difference value between request flow and response flow;

$$d[i] = f[i] - g[i] \quad (i = 0, 1 \ldots), \tag{14}$$

where $f[i]$ is a request flow and $g[i]$ is a response flow.

Usually, $d[i]$ is closely related to the network size, the number of hosts, and the sampling time. In order to counteract the influence of them, we convert it as follows:

$$K[i] = \begin{cases} 0 & i = 0 \\ \alpha * K[i-1] + (1-\alpha) * g[i] & i = 1, 2, \ldots \end{cases} \tag{15}$$

$$\text{DRRF}[i] = \frac{d[i]}{K[i]} \quad (i = 1, 2 \ldots). \tag{16}$$

In formula (15), $K[i]$ could be expressed as a recurrence relation of $g[i]$, where α is a custom constant, $\alpha \in [0..1]$. Thus, by using formula (16), we can get DRRF$[i]$, which will not be impacted by factors mentioned above. Instead, it is simply about current network traffic. We choose DRRF$[i]$ as another macroscopic feature to indicate the overall change of two-way communications caused by LDoS attacks.

So we can establish a component HMM based on DRRF feature, simply referred to as D-HMM.

4.3. Adjusting Threshold Dynamic. Enlightened by load-shedding method and Kaufman algorithm [33], we adjust the *threshold* value dynamically for improving the detection precision.

Let the $\Gamma[i]$ denote the mapping variable of the system effective payload and our algorithm threshold in the $(i+1)$th time span. Define $\Gamma[0] = 1$. The range of $\Gamma[i]$ values is in $[\Gamma[\min], 1]$, where $\Gamma[\min]$ is a rather small but not 0 constant. This is because if $\Gamma[\min]$ is 0, all data flows are not allowed to pass through it. Hypothesize that, right at the ith time over, the actual payload in the system is $\rho[i]$, and $\rho[\max]$ is the maximum number of payload, so we get $\varphi[i] = \rho[\max]/\rho[i]$. $\Gamma[i]$ could be presented in a recursive way as follows:

$$\Gamma[i] = \Gamma[i-1] * \varphi[i]. \tag{17}$$

And since $\Gamma[i] \in [\Gamma[\min], 1]$, we can get the final equation of $\Gamma[i]$; that is,

$$\Gamma[i] = \max\left\{\min\left\{\Gamma[0] * \prod_{j=1}^{i}\varphi[i], 1\right\}, \Gamma[\min]\right\}, \tag{18}$$

where $i = 1, \ldots, n$.

In this way, threshold value could be computed out by $\Gamma[i]$.

5. Experiments and Performance Results

In this section, we firstly introduce experimental environment setup. Then, we compare the normal flow with the attack one in aspect of the periodicity of WASPQ and the output of TF-HMM. Based on the comparisons, we validate the sufficient sensitivity of TF-HMM. Finally, we evaluate the performance results of TF-HMM in terms of detection rate, false-positive rate, average detection rate, training time, and recognition time.

5.1. Experimental Environment Setup. Data acquisition in real LDoS attacks is very difficult. Enlightened by papers [34–36], we construct experimental data by fusing controlled attack flows into real network background traffic.

To generate attack data, we have built a controlled experimental platform. 60 VMware hosts based on Windows XP system are chosen as user hosts. The collector and analyzer of network traffic are installed at Ubuntu 12.04 with Quad core 2.4 GHz CPU and 4 G RAM. We install Zombie tools at part of user hosts as bots. The controlled LDoS attack is launched by these bots, and then our experimental attack data could be achieved.

Accordingly, we choose a day's network traffic of a primary node in CERNET backbone networks as our experimental background traffic. There are 305985 records in the time window of 10 minutes. After the preprocessing, the background data contains 19877 hosts. Then, we fuse the attack data into the background traffic to evaluate TF-HMM performance.

5.2. Periodicity Analysis of WASPQ in P-HMM. The most obvious contrast between LDoS and normal flow is different periodicity in frequency domain. We firstly compare the normal WASPQ value with the attack WASPQ value.

As illustrated in Figure 2(a), in normal condition, the value of WASPQ is relatively high, almost 1100, because of the small proportion of short data packet in cache queue. In case of LDoS attacks, attackers use massive number of very short data packet to launch suddenly, and the value of WASPQ declines abruptly as shown in Figure 2(b), from about 1100 to 50. This is due to the fact that we use the weighted approach and highlight the importance of short packet in WASPQ calculation. We go on to draw the according periodograms of Figures 2(a) and 2(b). As you can see in Figure 2(d), in case of LDoS attacks, the change of WASPQ has obvious periodicity, while normal flow has none in Figure 2(c).

Next, we draw the corresponding PSD of WASPQ, as shown in Figure 3. We can see that there is a very wide frequency band in normal condition, but when attacking, the PSD value is almost below 51.5 Hz, and there is no distribution in higher frequency bands. We calculate the cumulative traffic spectrum (CTS) [5] of PSD, as shown in Figure 4. 98.65% power of attack flow distributes under 51.5 Hz. Relatively, 39.44% power of normal flow is lower than 51.5 Hz. The huge difference can make P-HMM the better detection sensitivity.

5.3. Comparison Output of TF-HMM in Normal and in Attack. In order to validate the sensibility of TF-HMM, we extract 30 seconds normal flow fragment firstly. Secondly, we extract 30 seconds fragment of LDoS violation and overlap them to one time axis. As shown in Figure 5, in normal, the value fluctuate in the range of −40~−984, while, under attacking, the peak value could reach 2.4~55 times more than normal value, or even larger. The red curve in Figure 5 obviously shows the 5 impulse low-rate violations, so it can be seen that TF-HMM has enough detection sensitivity to identify LDoS attacks hidden in legitimate network traffic.

5.4. Detection Rate and False-Positive Rate. In this section, we compare TF-HMM with representative nonparametric CUSUM algorithm [14] and PSD method [12] in detail. We focus on the detection accuracy and false positives of three algorithms in different network traffic. In order to evaluate impartially, various network traffics are employed in the following experiments, including different network utilization rates and attack intensions with or without legitimate periodicity flows. For simplicity, we call legitimate periodicity flows *the interference* in the sequel.

First, define detection rate R_d as

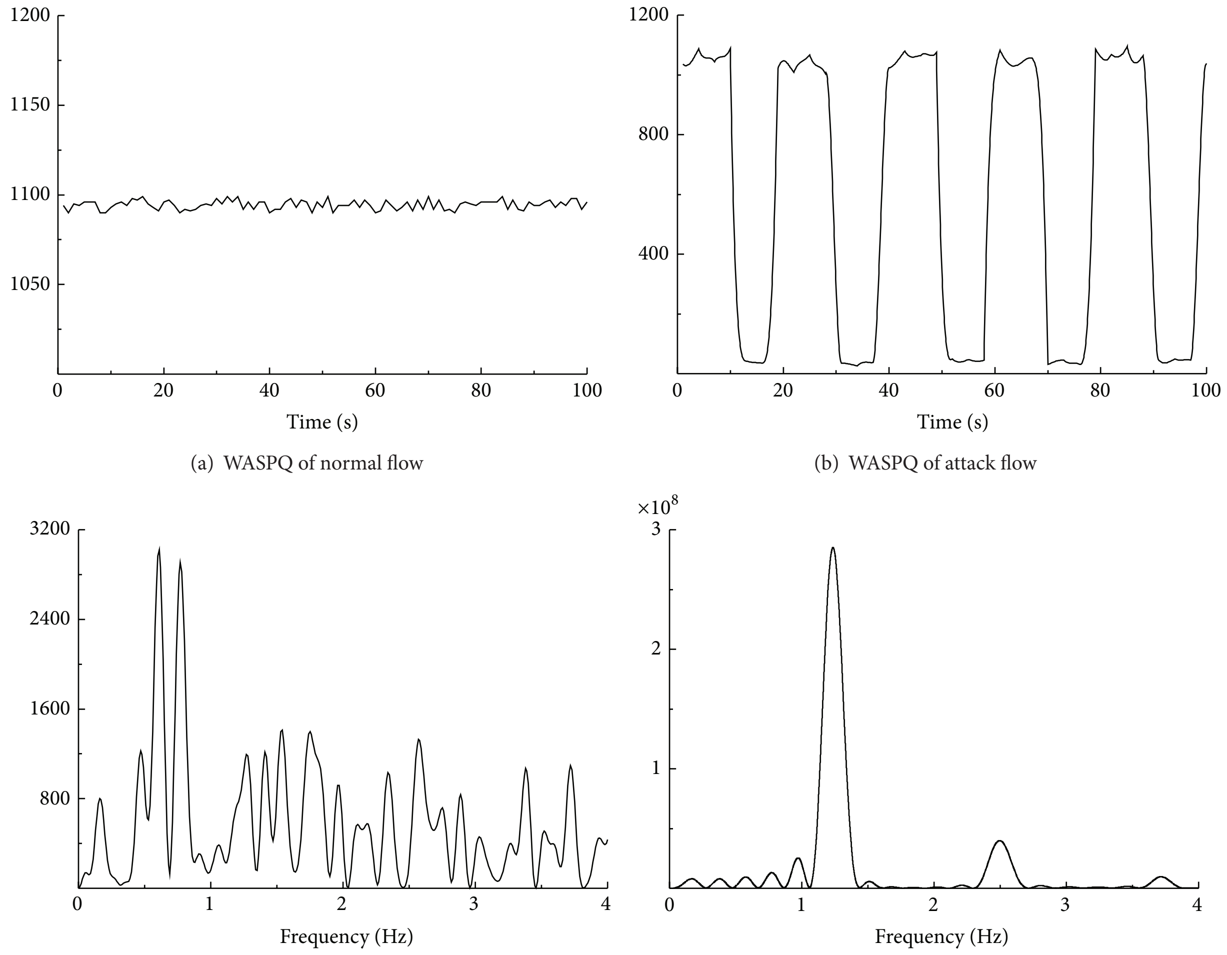
$$R_d = \frac{N_c}{N_r}. \tag{19}$$

Here, N_c is the number of attacks which have been detected correctly. N_r is the number of real attacks existing.

Next, define false-positive rate R_{fp} as

$$R_{\text{fp}} = \frac{N_a - N_c}{N_a}, \tag{20}$$

where N_a is the number of alarms by the detection algorithm and the difference between N_a and N_c is the number of false positives.



(a) WASPQ of normal flow

(b) WASPQ of attack flow

(c) Periodogram of normal flow

(d) Periodogram of attack flow

Figure 2: Comparison of WASPQ value and periodogram in normal and in attack flow.

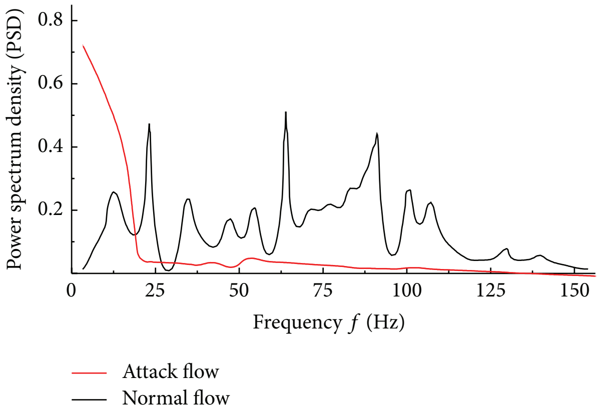


Figure 3: Comparison of normalized PSD of WASPQ in normal and in attack flow.

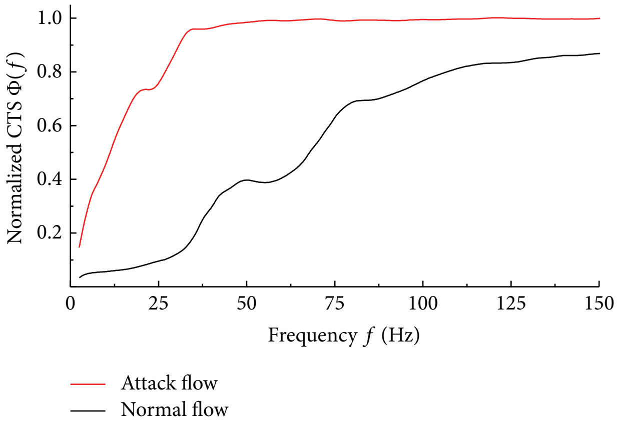


Figure 4: Comparison of CTS of WASPQ in normal and in attack flow.

The experiment results are shown as in Table 2.

(1) Without Attacks and without the Interference (See No. 1 Group). There are no periodicity flows, so periodicity-based algorithms (PSD and TF-HMM's P-HMM) give no false positives. However, CUSUM algorithm shows 3 false positives. This is because it is based on traffic volume accumulated

Table 2: Detection rate and false-positive rate comparison of 3 algorithms in different network traffic.

No.	Algorithm	N_a	N_c	N_r	Network utilization (%)	Interference
1	CUSUM	3	0	0	31.13	Without
	PSD	0	0			
	TF-HMM	0	0			
2	CUSUM	45	0	0	83.61	With
	PSD	18	0			
	TF-HMM	0	0			
3	CUSUM	23	1	2	33.15	Without
	PSD	2	2			
	TF-HMM	2	2			
4	CUSUM	45	19	30	47.23	With
	PSD	40	25			
	TF-HMM	33	29			
5	CUSUM	768	178	300	83.36	With
	PSD	588	250			
	TF-HMM	323	279			

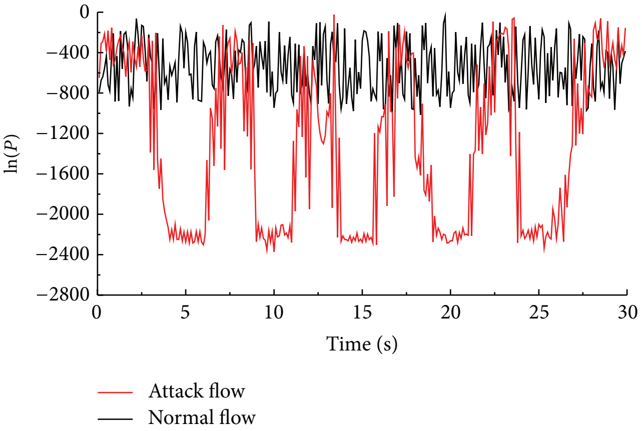


Figure 5: Comparison of output ln(P) of TF-HMM in normal and in attack flow.

method in the time domain, having no analysis capabilities of frequency domain.

(2) Without Attacks and with the Interference (See No. 2 Group). When injecting the interference flows and increasing utilization rate of network, false positives start appearing in the PSD algorithm but not in TF-HMM. This is due to the fact that the PSD cannot differentiate between the periodicity of the interference flows and one of pulse attacks, just capturing the periodicity. While TF-HMM's P-HMM can not only find the periodicity of flows but also analyze the WASPQ changes caused by LDoS attacks, it helps TF-HMM make an accurate distinction between legitimate periodicity flows and LDoS pulse flows.

Table 3: Average detection rate comparison of 3 algorithms.

CUSUM	PSD	TF-HMM
48.14%	69.39%	92.78%

(3) With Attacks and without the Interference (See No. 3 Group). Without the interference, the PSD and TF-HMM can identify exactly 2 times attacks hidden in background traffic based on the obvious periodicity of pulse attacks and show no false positives. But CUSUM still remains relatively high false positives because it is not a learning-oriented algorithm and is not also a frequency-domain-based one.

(4) With Attacks and with the Interference (See No. 4 and No. 5 Group). In No. 4 group, the result from TF-HMM is closer to REAL than other algorithms. Its R_{fp} is 12.12%. Conversely, the R_{fp} of CUSUM reaches up to 57.78%; the R_{fp} of PSD is 37.50%. With a growing intension of attacks and interferences, the R_{fp} of other two methods will be even higher.

In No. 5 group, we increased both of the attack intension and network utilization rate, and the advantages of TF-HMM based on multiple observed features becomes apparent. The R_{fp} of CUSUM is 76.82% and the R_{fp} of PSD is 57.48%, while its R_{fp} is 13.62%, far less than other algorithms. The reason for such low the R_{fp} of TF-HMM is that the two components of F-HMM and P-HMM play an important role.

When massive packets of legitimate periodicity flows and pulse attack flows arrive at the router, the PSD algorithm cannot accurately differentiate between them because it only uses the number of packet arrivals as a single periodic feature to find the periodicity in data sequence. Rather, the P-HMM can identify them because of WASPQ value abnormal decrease by pulse attacks (As illustrated in Figure 2(b)). Furthermore, the F-HMM can detect the packet's internal attribute fields that have been tampered with, because spoofed packets in pulse attacks result in abnormal fluctuations of WSFB.

The additive effect of combining multidimensional features starts to dominate, so we see a lower false-positive rate of TF-HMM. These provide some of the advantages of detection accuracy in TF-HMM not only with the higher detection rate, but also with the lower false-positive rate.

5.5. Average Detection Rate. In order to evaluate three detection approaches objectively, we varied attack intension, network utilization rate, sampling time, and the interference. Thus, there are obvious differences between every two groups. From the 100 groups of data gained, we calculated their average detection rate as presented in Table 3.

In Table 3, the average detection rate of PSD is obviously higher than CUSUM algorithm because it takes into account the inherent periodicity of LDoS violation. But the false positive rate is not still reduced to a reasonably low level; it limits the improvement of the detection accuracy. In contrast, since TF-HMM combines multiobserved features, its average detection rate reaches 92.78%, which is 1.93 times over CUSUM and 23.39% over PSD. It efficiently overcomes the bottleneck of limiting further increases in detection accuracy.

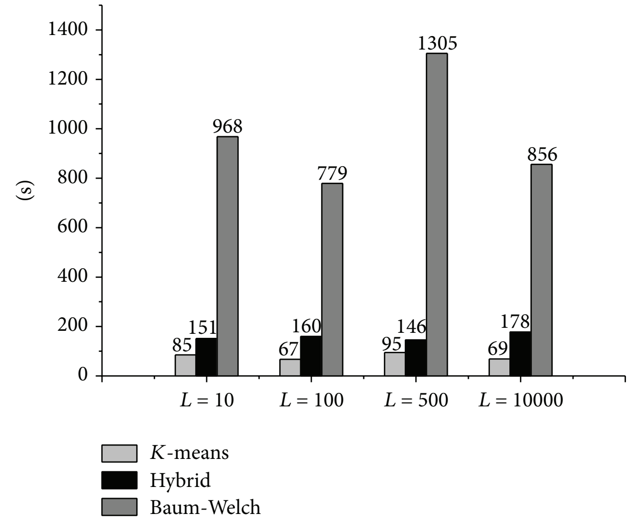


FIGURE 6: Influence of training time from training algorithm and segment length.

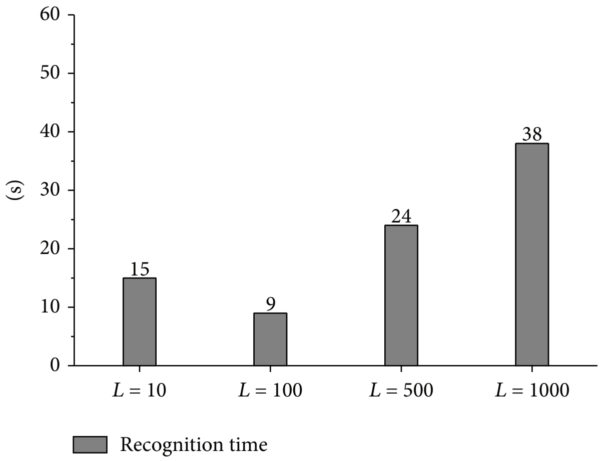


FIGURE 7: Influence of recognition time from segment length.

5.6. Training Time and Recognition Time. The time complexity of algorithms is vital to fast detection and response to QoS violation. Relevant experiments on training time and recognition time of the TF-HMM are sketched as in Figures 6 and 7.

As shown in Figure 6, the most time-consuming one is Baum-Welch algorithm; it is about 5 to 10 times of the other two algorithms; the second one is hybrid algorithm and then K-means algorithm. Furthermore, Baum-Welch algorithm is most sensitive to the length of segment. For example, using the same training sequence, the training time of $L = 500$ is 1.68 times more than the one of $L = 100$. But K-means and hybrid algorithms are insensitive to the length of segment.

And yet the recognition time of TF-HMM is short as shown in Figure 7. It is suitable for fast detection and responses to malicious QoS violations. Our ultimate goal is to achieve automated intrusion detection and responses in real time.

6. Conclusions

Current new LDoS violations are more and more characterized by high-distributed low rate. It is very difficult that fast detection and responses to stealthy LDoS streams are hidden in massive legitimate network traffic. The high false-positive rate is still the most striking bottleneck.

To overcome the bottleneck, our research contributions are summarized below in three technical aspects.

(1) Combining Multidimensional Features. Multiple micro- and macrofeatures, including WSFB, WASPQ, and DRRF, are combined together by using MF-HMM. The additive effects of combining multidimensional features make encouraging results on high detection rate with low false-positive rate.

(2) Synthesizing Methods in Frequency Domain and in Time Domain. Leveraging PSD analysis in the component P-HMM, we capture and identify the periodicity of LDoS pulse attacks in frequency domain. Furthermore, we calculate WSFB and DRRF feature in time domain by the components of F-HMM and D-HMM. These components make the accurate matching in detecting LDoS attacks at traffic streaming level.

(3) Adjusting Threshold Value Dynamically. Enlightened by load-shedding method and Kaufman algorithm, we adjust the threshold value dynamically to further reduce the false-positive rate.

For continued effort, we aim to improve the detection accuracy in complicated network traffic and ultimately to a fully automated process of detection and responses to LDoS attacks in real time.

Notations

CTS:	Cumulative traffic spectrum
CUSUM:	Cumulative Sum
D feature:	DRRF feature
DDoS:	Distributed denial of service
DFT:	Discrete Fourier transform
D-HMM:	The component HMM based on D feature
DoS:	Denial of service
DR:	Detection rate
DRRF:	Difference between request/response flows
DWT:	Discrete wavelet transform
F feature:	WSFB feature
F-HMM:	The component HMM based on F feature
LDoS:	Low-rate denial of Service
ln(P):	The output probability of TF-HMM
MF-HMM:	Multistream fused hidden Markov model
N-s:	Normal subsequence
P feature:	PSD of WASPQ feature
P-HMM:	The component HMM based on P feature
PSD:	Power spectrum density
QoS:	Quality of services
Q-s:	Questionable subsequence
RoQ:	Reduction of quality
TF-HMM:	Three-stream fused hidden Markov model
WASPQ:	Weighted average size of packet in queue
WSFB:	Weighted summation of flag bits.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to acknowledge the support of this work by National Natural Science Foundation of China (Grants nos. 60703023, 90204014) and Technology Development Plan of Jilin Province of China (Grant no. 20090110).

References

[1] A. Kuzmanovic and E. W. Knightly, "Low-rate TCP-targeted denial of service attacks: the shrew vs. the mice and elephants," in *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, Karlsruhe, Germany, 2003.

[2] X. Luo and R. K. C. Chang, "On a new class of pulsing denial-of-service attacks and the defense," in *Proceedings of the Network and Distributed System Security Symposium (NDSS '05)*, San Diego, Calif, USA, 2005.

[3] M. Guirguis, A. Bestavros, and I. Matta, "Exploiting the transients of adaptation for RoQ attacks on internet resources," in *Proceedings of the 12th IEEE International Conference on Network Protocols (ICNP '04)*, pp. 184–195, Berlin, Germany, October 2004.

[4] M. Guirguis, A. Bestavros, I. Matta, and Y. Zhang, "Reduction of Quality (RoQ) attacks on Internet end-systems," in *Proceedings of the IEEE International Conference on Computer Communication (INFOCOM '05)*, pp. 1362–1372, Miami, Fla, USA, March 2005.

[5] Y. Chen and K. Hwang, "Collaborative detection and filtering of shrew DDoS attacks using spectral analysis," *Journal of Parallel and Distributed Computing*, vol. 66, no. 9, pp. 1137–1151, 2006.

[6] G. Maciá-Fernández, J. E. Díaz-Verdejo, and P. García-Teodoro, "Mathematical model for low-rate dos attacks against application servers," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 3, pp. 519–529, 2009.

[7] Y. X. He, T. Liu, Q. Cao et al., "A survey of low-rate denial-of-service attacks," *Journal of Frontiers of Computer Science & Technology*, vol. 2, no. 1, pp. 1–17, 2008.

[8] Q. Zhu, Z. Yizhi, and X. Chuiyi, "Research and survey of low-rate denial of service attacks," in *Proceedings of the 13th International Conference on Advanced Communication Technology: Smart Service Innovation through Mobile Interactivity (ICACT '11)*, pp. 1195–1198, Gangwon-Do, Republic of Korea, February 2011.

[9] Z. J. Wu and B. S. Pei, "The detection of LDoS attack based on the model of small signal," *Acta Electronica Sinica*, vol. 39, no. 6, pp. 1456–1460, 2011.

[10] W. Zhi-Jun, Z. Hai-Tao, W. Ming-Hua, and P. Bao-Song, "MSABMS-based approach of detecting LDoS attack," *Computers & Security*, vol. 31, no. 4, pp. 402–417, 2012.

[11] Z.-J. Wu, H.-L. Zeng, and M. Yue, "Approach of detecting LDoS attack based on time window statistic," *Journal on Communications*, vol. 31, no. 12, pp. 55–62, 2010.

[12] C.-W. Zhang, J.-P. Yin, Z.-P. Cai, and W.-F. Chen, "RRED: robust RED algorithm to counter low-rate denial-of-service attacks," *IEEE Communications Letters*, vol. 14, no. 5, pp. 489–491, 2010.

[13] C. Yu, H. Kai, and Y.-K. Kwok, "Collaborative defense against periodic shrew DDoS attacks in frequency domain," *ACM Transactions on Information and System Security*, pp. 2–27, 2005.

[14] D. Liu, *Research on LDoS attack in soft-switch network [M.S. thesis]*, Communication and Information Engineering, 2013.

[15] A. Shevtekar, K. Anantharam, and N. Ansari, "Low rate TCP denial-of-service attack detection at edge routers," *IEEE Communications Letters*, vol. 9, no. 4, pp. 363–365, 2005.

[16] K. Chen, H. Y. Liu, and X. S. Chen, "Detecting LDoS attacks based on abnormal network traffic," *KSII Transactions on Internet and Information Systems*, vol. 6, no. 7, pp. 1831–1853, 2012.

[17] K. Chen, H. Liu, and X. Chen, "EBDT: a method for detecting LDoS attack," in *Proceedings of the IEEE International Conference on Information and Automation (ICIA '12)*, pp. 911–916, Shenyang, China, June 2012.

[18] D. Tang, K. Chen, X. Chen, H. Y. Liu, and X. Li, "Adaptive EWMA Method based on abnormal network traffic for LDoS attacks," *Mathematical Problems in Engineering*, vol. 2014, Article ID 496376, 11 pages, 2014.

[19] M. Yu, "An adaptive method for source-end detection of pulsing DoS attacks," *International Journal of Security and its Applications*, vol. 7, no. 5, pp. 279–288, 2013.

[20] Y. Xiang, K. Li, and W. Zhou, "Low-rate DDoS attacks detection and traceback by using new information metrics," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 2, pp. 426–437, 2011.

[21] Z. Zeng, J. Tu, B. Pianfetti et al., "Audio-visual affect recognition through Multi-stream Fused HMM for HCI," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, pp. 967–972, San Diego, Calif, USA, June 2005.

[22] H. Pan, S. E. Levinson, T. S. Huang, and Z.-P. Liang, "A fused hidden Markov model with application to bimodal speech processing," *IEEE Transactions on Signal Processing*, vol. 52, no. 3, pp. 573–581, 2004.

[23] M. Brand, N. Oliver, and A. Pentland, "Coupled hidden Markov models for complex action recognition," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 994–999, San Juan, Puerto Rico, June 1997.

[24] L. K. Saul and M. I. Jordan, "Mixed memory Markov models: decomposing complex stochastic processes as mixtures of simpler ones," *Machine Learning*, vol. 37, no. 1, pp. 75–87, 1999.

[25] L. R. Rabiner, "Tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[26] D. Zhou, H. Zhang, S. Zhang, and X. Hu, "DDoS attack detection method based on hidden Markov model," *Computer Research and Development*, vol. 42, no. 9, pp. 1594–1599, 2005.

[27] Z.-J. Wu and D. Zhang, "Attack simulation and signature extraction of low-rate DDoS," *Tongxin Xuebao/Journal on Communications*, vol. 29, no. 1, pp. 71–76, 2008.

[28] H.-P. Hu, J. Zhang, B. Liu, L. Chen, and X. Chen, "Simulation and analysis of distributed low-rate denial-of-service attacks," in *Proceedings of the 5th International Conference on Computer Sciences and Convergence Information Technology (ICCIT '10)*, pp. 620–626, IEEE, Seoul, Republic of Korea, December 2010.

[29] J. Zhang, H.-P. Hu, B. Liu, and F.-T. Xiao, "Detecting LDoS attack based on ASPQ," *Journal on Communications*, vol. 33, no. 5, pp. 79–84, 2012.

[30] P. D. Welch, "The use of the fast Fourier transform for estimation of spectra: a method based on time averaging over short, modified periodograms," *IEEE Transactions on Audio and Electroacoustics*, vol. 15, no. 2, pp. 70–74, 1967.

[31] H. C. So, Y. T. Chan, Q. Ma, and P. C. Ching, "Comparison of various periodograms for sinusoid detection and frequency estimation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 35, no. 3, pp. 945–952, 1999.

[32] J. Mirkovic and P. Reiher, "D-WARD: a source-end defense against flooding denial-of-service attacks," *IEEE Transactions on Dependable and Secure Computing*, vol. 2, no. 3, pp. 216–232, 2005.

[33] S. Kasera, J. Pinheiro, C. Loader, M. Karaul, A. Hari, and T. LaPorta, "Fast and robust signaling overload control," in *Proceedings of the 9th International Conference on Network Protocols (ICNP '01)*, pp. 323–331, Riverside, Calif, USA, November 2001.

[34] J. Francois, S. Wang, R. State et al., "BotTrack: tracking botnets using NetFlow and PageRank," in *NETWORKING 2011*, vol. 6640 of *Lecture Notes in Computer Science*, pp. 1–14, Springer, Berlin, Germany, 2011.

[35] H. L. Jiang, X. L. Shao, and Y. F. Li, "Online botnet detection algorithm using MapReduce," *Journal of Electronics and Information Technology*, vol. 35, no. 7, pp. 1732–1738, 2013.

[36] S. Nagaraja, P. Mittal, C. Hong et al., "BotGrep: finding P2P bots with structured graph analysis," in *Proceedings of the 19th USENIX Conference on Security*, Washington, DC, USA, 2010.