Modeling the Vulnerability of Feedback-Control Based Internet Services to Low-Rate DoS Attacks

Yajuan Tang, Xiapu Luo, Qing Hui, and Rocky K. C. Chang

***Abstract*— Feedback control is a critical element in many Internet services (e.g., quality-of-service aware applications). Recent research has demonstrated the vulnerability of some feedback-control based applications to low-rate denial-of-service (LRDoS) attacks, which send high-intensity requests in an ON/OFF pattern to degrade the victim's performance and evade the detection designed for traditional DoS attacks. However, the intricate interaction between LRDoS attacks and the feedback control mechanism remains largely unknown. In this paper, we address two fundamental questions: 1) what is the impact of an LRDoS attack on a general feedback-control based system and 2) how to conduct a systematic evaluation of the impact of an LRDoS attack on specific feedback-control based systems. To tackle these problems, we model the system under attack as a switched system and then examine its properties. We conduct the first theoretical investigation on the impact of the LRDoS attack on a general feedback control system. We formally show that the attack can make the system's steady-state error oscillate along with the attack period, and prove the existence of LRDoS attacks that can force the system to be far off the desired state. In addition, we propose a novel methodology to systematically characterize the impact of an LRDoS attack on specific systems, and apply it to a web server and an IBM Notes server. This investigation obtains many new insights, such as new attack scenarios, the bound of the system's states, the relationship between the bound and the LRDoS attacks, the close-formed equations for quantifying the impact, and so on. The extensive experimental results are congruent with the theoretical analysis.**

***Index Terms*— Feedback control, low-rate DoS attack, switched system, stability, performance degradation.**

I. Introduction

Feedback control is a fundamental building block of many Internet services. A classic example is the performance controller in a web server, which adjusts the server's configuration (e.g., admission rate) in response to the difference between the current and desired states for meeting expected performance (e.g., throughput and service time) [1]–[4]. Feedback control is also a central element in QoS-aware systems (e.g., cloud computing [5]–[7], high-performance computing [8], [9], virtualized servers [10], cyber-physical systems [11], and autonomic computing [12]).

Recent studies have demonstrated that Low-Rate DoS (LRDoS) attacks can degrade the performance of some feedback-control based applications (e.g., TCP [13]–[15], web server [16], [17], etc.). Different from flooding-based DoS attacks, LRDoS attacks send out intermittent (instead of continuous) high-volume requests to force the victim away from the desired state, thus deteriorating its performance. Moreover, LRDoS attacks can escape the detection designed for flooding-based DoS attacks because of their ON/OFF traffic patterns. A burst of requests can be termed an attack pulse, and an LRDoS attack then consists of a sequence of attack pulses. Although seminal studies have showed the possibility of using LRDoS to attack feedback-control based systems, and have studied the corresponding damage to a few applications under limited attack patterns [16], [17], they have not solved two fundamental questions. (1) What is the impact of an LRDoS attack on a general feedback control system? (2) How to conduct a systematic evaluation of the impact of an LRDoS attack on specific feedback-control based systems?

It is challenging to tackle these questions because LRDoS attacks can force a victim system to exhibit discontinuous behavior on the arrival of attack pulses. Traditional control theory cannot handle such situation because it targets a pure continuous (or discrete) system represented by differential (or difference) equations [18]. In this paper, to address the questions, we propose to model the system under attack as a switched system, which is a hybrid system composed of several subsystems and a switching law that indicates the sequence of subsystems [19].

We first model the impact of an LRDoS attack on a general feedback control system in two important respects (Section III): steady-state error and system state [20].

Manuscript received September 12, 2012; revised January 11, 2013 and September 6, 2013; accepted November 3, 2013. Date of publication November 20, 2013; date of current version January 28, 2014. This work was supported in part by the National Natural Science Foundation of China under Grants 60903185 and 61202396, in part by the Hong Kong GRF under Grant PolyU 5389/13E, in part by the Open Fund of Key Laboratory of Digital Signal and Image Processing of Guangdong Province, in part by the Academic Innovation Team Construction Project of Shantou University under Grant ITC12001, in part by the National Science Foundation under Grant 1241735, and in part by the Defense Threat Reduction Agency under Grants HDTRA1-10-1-0090 and HDTRA1-13-1-0048. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. T. Charles Clancy. *(Corresponding author: X. Luo.)*

Y. Tang is with the Department of Electronic and Information Engineering, Shantou University, Shantou 515000, China (e-mail: yjtang@stu.edu.cn).

X. Luo is with the Department of Computing and the Shenzhen Research Institute, The Hong Kong Polytechnic University, Hong Kong (e-mail: csxluo@comp.polyu.edu.hk).

Q. Hui is with the Department of Mechanical Engineering, Texas Tech University, Lubbock, TX 79409 USA (e-mail: qing.hui@ttu.edu).

R. K. C. Chang is with the Department of Computing, The Hong Kong Polytechnic University, Hong Kong (e-mail: csrchang@comp.polyu.edu.hk).

This paper has supplementary downloadable material at http://ieeexplore.ieee.org, provided by the author. The file contains additional theoretical analysis and experimental results. The material is 2.09 MB in size.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIFS.2013.2291970

We prove the existence of LRDoS attacks that can constrain the victim system to a state, which is determined by an attacker and diverges away from the desired state, by proving that the system under attack can still be Lyapunov and Lagrange stable. The investigation of the general feedback control system motivates us to develop a novel methodology to systematically analyze the impact of an LRDoS attack on specific feedback-control based systems.

We apply our methodology to two specific systems: the web server described in [16] and the feedback-control based IBM Notes server proposed in [21]. Although the web server has been studied in [16], our methodology empowers us to reveal many *new* insights in Section IV, such as, new attack scenarios, conditions under which an LRDoS attack will stabilize the web server, the relationship between the bound of the state of the web server and the LRDoS attacks, close-formed equations for characterizing how an LRDoS attack throttles the web server's admission rate, and the tradeoffs between the damage caused by an LRDoS attack and its cost. The IBM Notes server is different from the web server in terms of the application, the feedback controller, and the system model. To the best of our knowledge, we are the first to theoretically examine the impact of an LRDoS attack on such a server. Due to the page limit, we detail their proofs in the supplementary material.

In summary, our main contributions are as follows:

- To the best of our knowledge, we are the first to analytically investigate the impact of an LRDoS attack on a general feedback control system. We formally show that the system's steady-state error oscillates along with the attack. Moreover, we prove the existence of LRDoS attacks that can force the system to diverge away from the steady state, thus deteriorating its performance.
- We propose a novel methodology to systematically evaluate the impact of an LRDoS attack on specific feedback-control based systems. It models the system under attack as a switched system, identifies various attack scenarios and quantifies the impact of all scenarios. Moreover, it determines the conditions for LRDoS attacks to make the victim system Lyapunov and Lagrange stable, and identifies the bound of the system's state and the relationship between the bound and LRDoS attacks.
- We conduct the first comprehensive analysis on a feedback-control based web server under LRDoS attacks. We not only identify new attack scenarios that are *not* reported in previous work but also derive close-formed equations for quantifying the impact of all attack scenarios. Moreover, we prove conditions for the LRDoS attack to stabilize the web server, and obtain the bounds of the web server's state along with the relationship between the bound and LRDoS attacks.
- We carry out the first thorough analysis of the impact of LRDoS attacks on a feedback-control based IBM Notes server. By modeling the system under attack as a switched system, we identify various attack scenarios and quantify the impact of all attack scenarios. We also establish the conditions for the LRDoS attack to stabilize the IBM Notes server, and determine the bounds of the server's state and their relationship with the LRDoS attacks.
- We conduct extensive experiments through simulation and a testbed to evaluate the impact of LRDoS attacks. The results show that the LRDoS attack can cause severe damages regardless of whether the interval between consecutive attack pulses is fixed or randomized. We also demonstrate the tradeoff between the impact of the LRDoS attack and its cost, which indicates the existence of an optimal LRDoS attack.

The remainder of this paper is organized as follows. Section II reviews related work on LRDoS attacks. Section III investigates the effect of LRDoS attack on a general feedback control system. Section IV and Section V present a thorough examination on the damage caused by an LRDoS attack on a web server and an IBM Notes server, respectively. The experimental results are reported in Section VI. We discuss the practicality issues in Section VII and conclude this paper with future work in Section VIII.

II. Related Work

DDoS attacks have been plaguing the Internet for decades. They drain bandwidth and/or system resources to prevent normal users from receiving quality service [22]. Traditional flood-based attacks can be easily detected because of their continuously high sending rates [23], [24]. In contrast, low-rate DoS attacks have polymorphic traffic patterns and low average sending rates [13]–[15], [25].

A. Low-Rate DoS Attacks

LRDoS attacks was first proposed to throttle the throughput of TCP connections by causing intermittent packet losses [13]–[15]. Zhang et al. [26] and Schuchard et al. [27] showed that an attacker can launch LRDoS attacks on BGP sessions for crippling the Internets control plane. Recently, researchers examined the vulnerability of other applications to LRDoS attacks, including Internet services [16], [17], load balancers [28], wireless networks [29], and peer-to-peer networks [30].

Guirguis et al. found that an LRDoS attack can force a feedback control system to oscillate between the desired state and another state, and analyzed the effect of such attack on a web server [16], [17]. There are three major differences between our work and theirs. First, while Guirguis et al. described the possibility of launching the LRDoS attack on a feedback-control based system, we formally show that the LRDoS attack can compel a feedback control system to stay at a state other than the desired state by proving that the system under attack is *Lyapunov* and *Lagrange* stable. Second, we propose a novel methodology to systematically evaluate the impact of an LRDoS attack on specific systems. This methodology enables us to obtain many *new* insights that are not reported before. For example, for the same web server, we reveal in Section IV that an attacker can launch *three* types of LRDoS attacks by adjusting the attack period. Guirguis et al. only examined the third type of LRDoS attack [16]. It is worth noting that the other two types of attacks can cause *severer* damage to the web server, as shown in Fig. 4. Moreover, we thoroughly analyze each type of LDRoS attack, including giving closed-form expressions for the throttled admission rate, determining the conditions under which the LRDoS attack will make the web server Lyapunov and Lagrange stable, deciding the bound of the system's state and the relationship between the bound and the LRDoS attack, and identifying the relationship between the damage caused by an LRDoS attack and its cost. Third, we employ our methodology to analyze the vulnerability of a feedback-control based IBM Notes server to the LRDoS attack. Note that the IBM Notes server is different from the web server in three aspects, including the application (i.e., email service vs. web service), the feedback controller (i.e., I controller vs. PI controller), and the system model. Our analysis formally shows that an LRDoS attack can cause severe damage to the IBM Notes server. The investigation of these two kinds of servers provides convincing evidences on the threat of LRDoS attacks to feedback-control based systems and the generality of our methodology.

Maciá-Fernández et al. proposed a smart attack called Low Rate DoS attack against Application Servers (LoRDAS), which dispatches attack requests to the victim server at carefully selected instances to occupy the server's queue and consequently prevent it from serving legitimate requests [31], [32]. They also built a mathematical model for LoRDAS attacks and evaluated their performance. The major difference between our work and theirs is that the attack examined in our paper exploits the feedback control mechanism in the victims system. However, LoRDAS attack takes advantage of the victim's queue. Moreover, an LoRDAS attack needs to predict the time instants when the queue of the victim server has free position so that it can make the attack requests reach the server around these time instants [33], whereas an LRDoS attack can send attack pulses with fixed or randomized interval to the victim system.

B. Defending Against LRDoS

As LRDoS attacks have ON/OFF traffic patterns, they can evade detection schemes targeting flooding-based DoS attacks and therefore have motivated the design of new detection approaches [15], [25], [34]–[39]. However, these approaches cannot be directly used to detect LRDoS attacks against Internet services for two reasons. First, as all of these approaches aim at LRDoS attacks targeting TCP or other systems (e.g., wireless networks, P2P networks, etc.), they rely on features specific to TCP and those systems. For example, we proposed the detection of anomalies in incoming TCP data traffic and outgoing TCP ACK traffic [15]. Shevtekar et al. regarded a TCP flow as malicious if its period is equal to the fixed minimal RTO and its burst length is no less than other connections' RTTs [36]. To detect LRDoS attacks using spoofed IP addresses, Shevtekar et al. captured anomalies that short-lived flows occupy a high percentage of the total traffic going through a link [40]. We proposed a new metric named the congestion participation rate (CPR) to infer attack flows that try to send more packets during congestion [41]. To detect distributed LRDoS attacks, Xiang et al. [39] used generalized entropy and information distance to quantify anomalies in packets, and required the control of all routers in the network. However, the detection of LRDoS attacks aimed at Internet services requires new metrics [42].

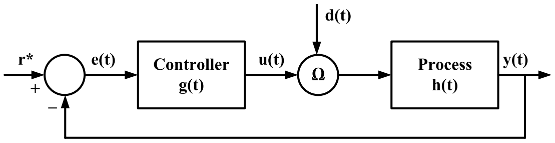


Fig. 1. A general feedback control system.

Second, the majority of the previous work focuses on the Shrew attack [13] that has a fixed attack period equal to TCP's minimal RTO. For example, the spectral-analysis approach relies on the traffic spectrum of Shrew attack flows, which is different from that of normal flows because of the even attack period [35], [38]. However, LRDoS attacks can change their attack periods for mimicking normal flows. Sun et al. suggested using autocorrelation and dynamic time warping (DTW) to detect Shrew attacks, because their traffic bursts are the same and have fixed periods [34]. However, it is unnecessary for LRDoS attacks to have invariable periods and similar attack pulses [25].

III. Attacking a General Feedback Control System

In this section, we investigate the impact of an LRDoS attack on a general feedback control system as shown in Fig. 1. This system comprises two major components: a *process* (i.e., $h(t)$) and a *controller* (i.e., $g(t)$). $h(t)$ represents any Internet service (e.g., web service, video streaming, etc.) while $g(t)$ generates a *control signal* (i.e., $u(t)$) to regulate $h(t)$ [2]. The input to the controller is a *control error* (i.e., $e(t)$), which is the difference between $h(t)$'s output (i.e., $y(t)$) and the expected value r^*. $y(t)$ can be any measurable metric, such as system utilization or queue length. r^* is usually selected for the system to achieve the best performance by the system designer, and the controller drives $y(t)$ towards r^* based on $e(t)$. $d(t)$ denotes the arrival rate of requests. To simplify the discussion, we assume that the arrival rate of normal requests is constant, denoted as λ_n. We compare the results when $d(t)$ is constant or random for two real systems in Section VI. Depending on the application, $d(t)$ enters into the process through an additive or a multiplicative operator Ω [16], [43]. The steady-state error, defined as $e_s(t) = \lim_{t\to\infty}(e(t))$, is the first performance index of a feedback control system [18]. As $e_s(t)$ quantifies the accuracy of control, it should be as small as possible and preferably zero.

Since all practical feedback control systems are necessarily stable [18], we assume that the victim system is Lyapunov stable without attack, meaning that the system trajectory can be kept arbitrarily close to an equilibrium point by starting sufficiently close to it [44]. We also assume that the control error, the output, and the control signal induced by an attack pulse are not equal to zero. Otherwise, the LRDoS attack cannot have any effect on this system.

We examine the impact of LRDoS attacks on a general feedback control system from *two* perspectives. First, we formally show in Proposition 1 that an LRDoS attack makes the *steady-state error* (i.e., $e_s(t)$) oscillate according to the attack pattern instead of converging to zero (Section III-A). Second, by modeling the system under an LRDoS attack as a *switched system* [19], [45], we prove the existence of LRDoS attacks that can force the victim system to diverge away from the desired state and constraint it to a state determined by the attacker through Proposition 2-3. To the best of our knowledge, we are the first to theoretically reveal how an LRDoS attack degrade the performance of feedback-control based systems.

The insights obtained in this section motivate us to propose a novel methodology to systematically analyze the impact of an LRDoS attack on a specific system. It includes four steps: (1) model the victim system under attack as a switched system; (2) quantify the impact on each subsystem and identify various attack scenarios; (3) determine the condition for an LRDoS attack to make the victim system Lyapunov stable and Lagrange stable; (4) decide the bound of the victim system's state and the relationship between the bound and the LRDoS attack. We apply this methodology to examine a web server and a feedback-control based IBM Notes server.

A. Steady-State Error

An LRDoS attack transmits intermittent attack pulses to a target system. Let τ_k, $k \in \mathbb{Z}^+$, represent the interval between the k^{th} and $(k+1)^{th}$ attack pulses, during which the LRDoS attack sends nothing. Note that the attack is not necessarily periodic (i.e., τ_{k1} is not necessarily equal to τ_{k2}, $k1 \neq k2$). For simplifying the ensuing discussion, we assume that the requests in each attack pulse have the same arrival rate, denoted as λ_a. Then, we can model the LRDoS attack as a sequence of Dirac signals: $\sum_{k=1}^{\infty} \lambda_a \delta(t - T_k)$, where T_k is the arrival time of the k^{th} pulse (i.e., $T_{k+1} - T_k = \tau_k$). Consequently, we have $d(t) = \begin{cases} \lambda_n & t \neq T_k \\ \lambda_a + \lambda_n & t = T_k. \end{cases}$

We use capital letters to denote the Laplace transform of a component (i.e., $G(s)$ and $H(s)$) and the input (i.e., $D(s)$), and employ $*$ to denote the convolution operator. According to Fig. 1, we have

$$y_a(t) = \begin{cases} r^* + \sum_k \lambda_a u(T_k) \mathcal{L}^{-1}[\frac{H(s)}{1-\lambda_n G(s) H(s)}] \delta(t - T_k), \\ \hfill \text{additive } \Omega, \\ r^* + \lambda_a \sum_k u(T_k) h(t - T_k), \\ \hfill \text{multiplicative } \Omega, \end{cases}$$

$$u_a(t) = \begin{cases} -\sum_k \lambda_a u(T_k) \mathcal{L}^{-1}[\frac{H(s)}{1-\lambda_n G(s) H(s)}] \delta(t - T_k) * g(t), \\ \hfill \text{additive } \Omega, \\ -\lambda_a \sum_k u(T_k) h(t - T_k) * g(t), \\ \hfill \text{multiplicative } \Omega, \end{cases}$$

Although a feedback control system aims at minimizing the control error in the steady state, Proposition 1 shows that the steady-state error $e_s(t) (= \lim_{t \to \infty} e(t))$ cannot become zero (i.e., the system cannot stay at or converge to the desired state). More precisely, under an LRDoS attack, the control error oscillates with the attack period, and its magnitude is affected by λ_a. Hence, an attacker can cause different levels of damage by varying the attack period and λ_a.

Proposition 1: In the presence of an LRDoS attack, the steady-state error $e_s(t)$, steady-state output $y_s(t)$, and steady-state control signal $u_s(t)$ oscillate according to the attack period τ_k, $k \in \mathbb{Z}^+$ and λ_a.

B. Switched System Model

Let $x(t)$ be the system state consisting of $y(t)$ and $u(t)$ [44]. A dynamic system is represented by $\dot{x} = f(x)$ and a solution of this equation corresponds to a curve in the state space that is also referred to as a system trajectory [44]. We use $x(t) = x_z(t) + x_n(t)$ and $x'(t) = x_z(t) + x_n(t) + x_a(t)$ to denote system states in the absence/presence of an LRDoS attack, respectively. x_z, x_n, and x_a are system states caused by zero input, normal requests, and attack requests, individually.

Compared with $x(t)$, the extra component $x_a(t)$ in $x'(t)$ denotes discrete events in the system model $f(x)$, because its components (i.e., $y_a(t)$ and $u_a(t)$) appear at T_k, $k \in \mathbb{Z}^+$. A continuous-time system with discrete switching events is referred to as a switched system, which consists of a family of continuous-time subsystems and a switching rule that governs the switching between them [19], [45].

We first consider a family of subsystems given by $\dot{x} = f_p(x)$, where $p \in \mathcal{P}$, $\mathcal{P}$ is a finite index set, and for each $p \in \mathcal{P}$, f_p is Lipschitz continuous. A switched system consists of a sequence of these subsystems

$$\dot{x} = f_\sigma(x), \tag{1}$$

where $\sigma(t) : [0, \infty) \to \mathcal{P}$ is an index of the active subsystem. When $t \in [t_k, t_{k+1})$, $\sigma(t) = i_k$, $k, i \in \mathbb{Z}^+$ (i.e., the i_k^{th} subsystem is active in $t \in [t_k, t_{k+1})$). Hence, $\sigma(t)$ is a piecewise constant. At time instant t_{k+1}, $\sigma(t)$ changes from i_k to i_{k+1}, and therefore we call t_k, $k \in \mathbb{Z}^+$, the switching times. The state $x(t)$ of the switched system (1) is defined as the state $x_{i_k}(t)$ of the i_k^{th} subsystem when $t \in [t_k, t_{k+1})$.

There are two types of switching points: time-dependent switching point and state-dependent switching point. The former is determined by attack pulses (i.e., switching happens at T_k). The latter is caused by the system (i.e., s_i, $i \in \mathbb{Z}^+$). We use t_k to denote all switching points (i.e., $t_k = \{T_k, s_k\}$). For ensuring the causality of the switching times (i.e., $t_{k+1} > t_k > 0$), we assume that if there are an infinite number of switching times, there exists $\vartheta > 0$ such that for every constant $T \geq 0$ one can find a positive integer k for which $t_{k+1} - \vartheta \geq t_k \geq T$.

C. Stability Analysis

Without loss of generality, we assume that the switched system has multiple equilibrium points, each of which is a state x_e such that $f(x_e) = 0$ [44], [46]. Some of the equilibrium points result from the intrinsic feedback control mechanism while others may be caused by the LRDoS attack. Note that the steady state (i.e., the expected state) is one of the equilibrium points [47]. We discuss the case when the switched system has only one equilibrium point in the supplementary material.

We first prove in Proposition 2 the existence of LRDoS attacks that can make the switched system (i.e., the system under attack) Lyapunov stable by using the *multiple Lyapunov* *function* approach [48]. In other words, this finding indicates the existence of LRDoS attacks that can cause the victim system to stay at an equilibrium point other than the steady state, because a Lyapunov stable switched system is stable at any equilibrium point [48].

Besides Lyapunov stability, we also consider Lagrange stability that refers to the stability of the trajectory, not the stability of equilibrium points. For a Lagrange stable system, its system state is bounded for every initial condition in a neighborhood [20]. We prove in Proposition 3 that the switched system's trajectory is bounded and there exist LRDoS attacks that can restrict the system trajectory to a level determined by the attacker as shown in Eqn. (2).

Proposition 2: Given that the victim system is Lyapunov stable in the absence of LRDoS attacks, for each $p \in \mathcal{P}$ and every pair of switching times $t_i < t_j$ where $\sigma(t_i) = \sigma(t_j) = p$, if there exists an attack sequence $\{t_k\}$ such that for all $i, j \in \mathbb{Z}^+$, $t_j - t_i \geq \frac{N_e \kappa_e}{\varepsilon}$, where $\varepsilon > 0$, N_e is the number of subsystems between $t \in [t_i, t_j)$, and $\frac{V_{p1}}{V_{p2}} \leq \kappa_e$, $\kappa_e > 0$ for any $p_1 \neq p_2$, then the victim system is Lyapunov stable in the presence of LRDoS attack.

Proposition 3: Given that the victim system is Lyapunov stable in the absence of LRDoS attacks, the victim system is Lagrange stable in the presence of LRDoS attacks. Furthermore, if there exists an attack sequence $\{t_k\}$ such that for all $i, j \in \mathbb{Z}^+$, $t_j - t_i \geq \frac{N_e \kappa_e}{\varepsilon}$, where $\varepsilon > 0$, N_e is the number of subsystems between $t \in [t_i, t_j)$, and $\frac{V_{p1}}{V_{p2}} \leq \kappa_e$, $\kappa_e > 0$ for any $p_1 \neq p_2$, then

$$x(t) \leq (\kappa_e^{N_e} e^{-\varepsilon N_e \Delta T})^{\frac{1}{2}} |x(0) - x_{ep}| + |x_{ep}|. \tag{2}$$

IV. Attacking a Web Server

Besides the qualitative analysis of the impact of an LRDoS attack on a general feedback control system, we further *quantify* it through real systems. As web servers are becoming the major platform for providing Internet services, we conduct a comprehensive investigation into the impact of an LRDoS attack on the web server described in [16]. As another example, we investigate the impact of an LRDoS attack on an IBM Notes server proposed in [21] (Section V). It is worth noting that our methodology can be applied to other feedback-control based Internet services.

For the web server, we address three challenging questions that were *not* considered in [16].

1. Are there other types of LRDoS attacks besides the one studied in [16]? If yes, what are they?

The stability analysis in Section III reveals that there are other types of LRDoS attacks in addition to the one examined in [16], because the attack in [16] allows the system to return to the steady state whereas Propositions 2-3 in Section III-C prove that an LRDoS attack can force the system to stay at an equilibrium point determined by the attacker. Motivated by this insight, we identify *three* types of LRDoS attacks in Section IV-B. It turns out that the attack in [16] is one type of LRDoS attacks, which is less severe than the other types in terms of the damage incurred to the web server.

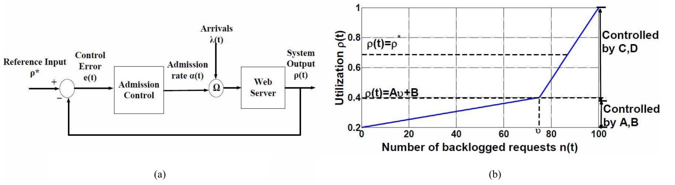


Fig. 2. A web server and its parameter relationship. (a) A feedback control based web server. (b) The relationship between utilization and the number of backlogged requests. Here the parameters are from [16] and $1 \geq \rho^* \geq \mathrm{A}\upsilon + \mathrm{B}$.

2. If the answer to the first question is true, what are the impact and the effectiveness of these attacks?

As Section IV-B uncovers new LRDoS attacks, we quantify their impact on the web server. It is non-trivial to model the impact of these new attacks compared to the one in [16], because for the new attacks we have to determine the system state immediately before a new attack pulse arrives. For the attack in [16], the system is in the steady state before a new attack pulse reaches it. Motivated by Proposition 1 for the general feedback control system, we prove in Proposition 4 (Section IV-C) that a periodic LRDoS attack will cause the web server's state to converge with a periodic solution. Moreover, Proposition 5 gives closed-form equations for the maximal and minimal values of the admission rate constrained by different types of LRDoS attacks. Beside examining the impact, we also investigate the effectiveness of the LRDoS attacks by defining two metrics and modeling the relationship between the metrics and the parameters of an LRDoS attack (Proposition 6-7). The result implies the existence of optimal attack patterns, which will be studied in future work.

3. What kind of LRDoS attacks can make the web server under attack Lyapunov and Lagrange stable? What is the bound of the state of the web server under attack?

Since Proposition 2 only proves the existence of such LRDoS attacks, we determine the conditions for an LRDoS attack in Proposition 8 and characterize the relationship between the bound of the web server's state and the parameters of an LRDoS attack in Proposition 9.

A. The Web Server Model

Fig. 2(a) illustrates the web server model. It employs a Proportional-Integral (PI) controller to adjust the admission rate (i.e., $\alpha(t)$) according to the difference between the desired utilization (i.e., ρ^*) and the actual utilization (i.e., $\rho(t)$), which is affected by the number of backlogged requests. Therefore, the system state can be described by the admission rate $\alpha(t)$, the utilization $\rho(t)$, and the number of backlogged requests $n(t)$, as follows:

$$\dot{\alpha}(t) = K(\rho^* - \rho(t)), \ \alpha(t) \in [0, 1],$$

$$\rho(t) = \begin{cases} An(t) + B & \text{if } n(t) < \upsilon, \\ Cn(t) + D & \text{if } \frac{1-D}{C} \geq n(t) \geq \upsilon, \\ 1 & \text{if } n(t) > \frac{1-D}{C}, \end{cases} , \rho(t) \in [0, 1], \tag{3}$$

$$\dot{n}(t) = \lambda\alpha(t) - \mu^{\mathcal{W}}, \ n(t) \in [0, +\infty), \tag{4}$$

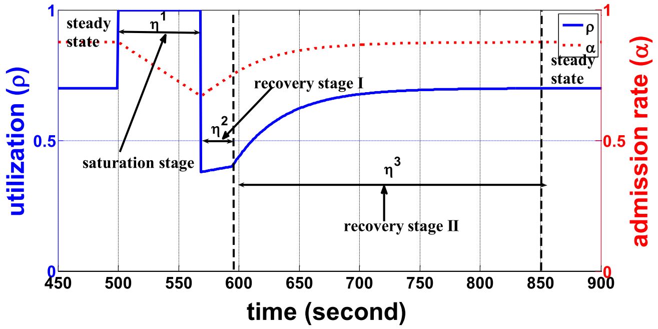


Fig. 3. The effect of one attack pulse at $t = 500$ on the admission rate and utilization.

where $\mu^{\mathcal{W}}$ is the service rate and $\rho(t)$ is a piecewise function with constants A, B, C, D, and v.

Note that this system is the same as the one in [16], except that we adopt a continuous-time model because it is more realistic and assume that $\mu^{\mathcal{W}}$ is constant for analytical tractability. Following [16], we assume that the arrival rate of normal requests is a constant $\lambda_n^{\mathcal{W}}$. We evaluate the effects of stochastic arrival processes in Section VI-A and the supplementary material. Similar to Section III, we assume the arrival rate of attack requests in each attack pulse is a constant $\lambda_a^{\mathcal{W}}$. We assume that the desired utilization ρ^* lies in the range of $[Av+B, 1]$. It is easy to extend the result to the scenario when ρ^* is between 0 and $Av+B$. Fig. 2(b) shows the relationship between $\rho(t)$, ρ^*, and $n(t)$, where $1 \geq \rho^* \geq Av + B$.

B. Different Types of the LRDoS Attack

The goal of an LRDoS attack is to throttle the web server's admission rate so that requests from normal users are dropped. To achieve this, an attacker sends intermittent attack pulses to cause transient congestion in the web server, which forces the server to decrease its admission rate.

On the arrival of an attack pulse, the server moves through three different stages before returning to the steady state: *saturation*, *recovery I*, and *recovery II*, as shown in Fig. 3. We use $\eta^{\mathcal{W},1}$, $\eta^{\mathcal{W},2}$, and $\eta^{\mathcal{W},3}$ to denote the durations of these three stages. The saturation stage begins right after the arrival of an attack pulse. During this stage, the utilization equals 1 (i.e., $\rho(t) = 1$) and the admission rate decreases. After $\rho(t) < \rho^*$, the server enters two recovery stages consecutively, during which the admission rate and the utilization restore to the steady state. The difference between the two recovery stages lies in the model for $\rho(t)$ and $n(t)$ in Eqn. (3).

Saturation stage: Once an attack pulse arrives, the system enters the saturation stage with $\rho(t) = 1$. The system state during this stage is characterized by $\rho(t) = 1$, $\dot{\alpha}(t) = K(\rho^* - 1)$, and $n(t) = \frac{1}{2}\lambda_n^{\mathcal{W}} K(\rho^*-1)t^2 + (\lambda_n^{\mathcal{W}}\alpha_0 - \mu^{\mathcal{W}})t + (\lambda_a^{\mathcal{W}} + \lambda_n^{\mathcal{W}})\alpha_0$, where α_0 is the initial value of $\alpha(t)$ and the expression for $n(t)$ is obtained by substituting $\alpha(t)$ in Eqn. (4) and then solving the differential equation. When the number of backlogged requests is reduced to $n(t) = (\rho^* - D)/C$ if $\rho^* \geq Av + B$, we have $\rho(t) \leq \rho^*$ and this stage ends. $\eta^{\mathcal{W},1}$ can be obtained by solving $n(\eta^{\mathcal{W},1}) = (\rho^* - D)/C$ with the initial conditions $[\alpha_0, \rho_0, n_0]$, where $\alpha_0 = \alpha(0^-) = \alpha(0^+)$, $n_0^+ = (\lambda_n^{\mathcal{W}} + \lambda_a^{\mathcal{W}})\alpha_0 + n_0^-$, and $\rho_0 = \rho(0^+) = 1$: at the bottom of this page.

Recovery stage I: At the beginning of this stage, because $\rho(\eta^{\mathcal{W},1}) \leq \rho^*$, $\alpha(t)$ stops decreasing and begins increasing. Consequently, $\rho(t)$ also increases. The initial conditions for this stage include $n(\eta^{\mathcal{W},1-}) = n(\eta^{\mathcal{W},1+}) = \lambda_n^{\mathcal{W}}\alpha(\eta^{\mathcal{W},1+})$, $\alpha(\eta^{\mathcal{W},1-}) = \alpha(\eta^{\mathcal{W},1+}) = \alpha_0 + K(\rho^* - 1)\eta^{\mathcal{W},1}$, and $\rho(\eta^{\mathcal{W},1+}) = A\lambda_n^{\mathcal{W}}\alpha(\eta^{\mathcal{W},1+}) + B$. The evolution of the system state is given by $\rho(t) = A\lambda_n^{\mathcal{W}}\alpha(t) + B$, $\dot{\alpha}(t) = K(\rho^* - \rho(t))$ and $\dot{n}(t) = \lambda_n^{\mathcal{W}}\alpha(t) - \mu^{\mathcal{W}}$. This stage ends when $n(t) = v$, and then the system enters the recovery stage II. $\eta^{\mathcal{W},2}$ can be obtained by solving $\rho(\eta^{\mathcal{W},2}) = Av + B$ with the initial conditions $\alpha(\eta^{\mathcal{W},1+})$, $\rho(\eta^{\mathcal{W},1+})$, and $n(\eta^{\mathcal{W},1+})$:

$$\eta^{\mathcal{W},2} = \frac{1}{A\lambda_n^{\mathcal{W}}K} \ln \frac{A\lambda_n^{\mathcal{W}}\alpha(\eta^{\mathcal{W},1}) + B - \rho^*}{Av + B - \rho^*}.$$

Recovery stage II: The differences between recovery stages I and II lie in the parameters and the initial conditions. The initial conditions here are $\alpha(\eta^{\mathcal{W},2-}) = \alpha(\eta^{\mathcal{W},2+}) = \frac{v}{\lambda}$, $\rho(\eta^{\mathcal{W},2-}) = \rho(\eta^{\mathcal{W},2+}) = Av + B$, and $n(\eta^{\mathcal{W},2-}) = \dot{n}(\eta^{\mathcal{W},2+}) = \lambda_n^{\mathcal{W}}\alpha(\eta^{\mathcal{W},2+}) - \mu^{\mathcal{W}}$. This stage ends when the utilization reaches the desired value. Thus, $\eta^{\mathcal{W},3}$ can be obtained by solving $\rho(\eta^{\mathcal{W},3}) = \rho^*$ with the initial conditions $\alpha(\eta^{\mathcal{W},2+})$, $\rho(\eta^{\mathcal{W},2+})$, and $n(\eta^{\mathcal{W},2})$:

$$\eta^{\mathcal{W},3} = \frac{1}{C\lambda_n^{\mathcal{W}}K} \ln \frac{C\lambda_n^{\mathcal{W}}\alpha(\eta^{\mathcal{W},2}) + D - \rho^*}{b\rho^* - \rho^*},$$

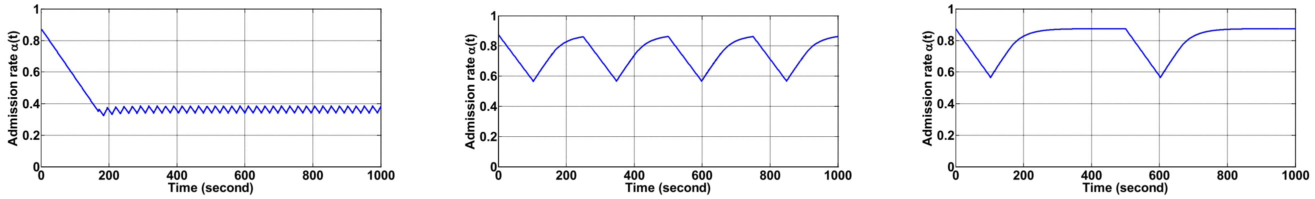
where $b \approx 1$ and $\alpha(\eta^{\mathcal{W},2}) = \frac{v}{\lambda}$.

According to the relationship between the attack period and the duration of the three stages, we identify *three* types of LRDoS attacks that have different impacts on the web server. Fig. 4 demonstrates the admission rate's trajectory in the presence of the different LRDoS attacks.

Type I attack: It has $\tau_{k+1} < \eta_k^{\mathcal{W},1} + \eta_k^{\mathcal{W},2}$. Under such an attack, the admission rate's trajectory involves two stages: the saturation stage and the recovery stage I, as illustrated in Fig. 4. When $\eta_k^{\mathcal{W},1} < \tau_{k+1} < \eta_k^{\mathcal{W},1} + \eta_k^{\mathcal{W},2}$, a new attack pulse arrives during recovery stage I.

When $\tau_{k+1} \leq \eta_k^{\mathcal{W},1}$, a new attack pulse reaches during the saturation stage. In this case, the system state oscillates between the saturation stage and the recovery stage I for $k < k_0$, where k_0 is a finite constant. The reasons behind such behavior are as follows. The $(k+1)^{th}$ attack pulse arrives when the system is still saturated due to the requests in the k^{th} attack pulse. Therefore, the admission rate continues decreasing and the number of backlogged requests also decreases because less requests are admitted. If $\tau_{k+1} \leq \eta_k^{\mathcal{W},1}$ holds for all the attack pulses, $\alpha(t)$ will be kept at zero. In the extreme case, it

$$\eta^{\mathcal{W},1} = \frac{(\lambda_n^{\mathcal{W}}\alpha_0 - \mu^{\mathcal{W}})}{\lambda_n^{\mathcal{W}}K(1-\rho^*)} - \frac{\sqrt{(\lambda_n^{\mathcal{W}}\alpha_0 - \mu^{\mathcal{W}})^2 - 2\lambda_n^{\mathcal{W}}K(\rho^* - 1)((\lambda_n^{\mathcal{W}} + \lambda_a^{\mathcal{W}})\alpha_0 + n_0 - \frac{\rho^*-D}{C})}}{\lambda_n^{\mathcal{W}}K(1-\rho^*)}$$



(a) (b) (c)

Fig. 4. Three types of LRDoS attacks and their impact on the admission rate. (a) Type I attack. (b) Type II attack. (c) Type III attack.

becomes the flooding-based DoS attack. Such kind of attack is not efficient because the majority of the attack requests will be dropped. Therefore, we ignore this situation and redefine the type I attack as $\eta_k^{\mathcal{W},1} < \tau_{k+1} < \eta_k^{\mathcal{W},1} + \eta_k^{\mathcal{W},2}$.

Type II attack: It has $\eta_k^{\mathcal{W},1} + \eta_k^{\mathcal{W},2} \leq \tau_{k+1} < \eta_k^{\mathcal{W},1} + \eta_k^{\mathcal{W},2} + \eta_k^{\mathcal{W},3}$. Then, the $(k+1)^{th}$ attack pulse arrives when the system is in the recovery stage II. In this case, the trajectory involves three stages: the saturation stage and the recovery stages I and II, as shown in Fig. 4. However, the admission rate cannot restore to α^c, which is the desired value when $\rho = \rho^*$, because $\tau_{k+1} < \eta_k^{\mathcal{W},1} + \eta_k^{\mathcal{W},2} + \eta_k^{\mathcal{W},3}$.

Type III attack: It has $\tau_{k+1} \geq \eta_k^{\mathcal{W},1} + \eta_k^{\mathcal{W},2} + \eta_k^{\mathcal{W},3}$. Then the evolution of the system involves all four stages: the saturation stage, the recovery stage I, the recovery stage II, and the steady state, as illustrated in Fig. 4. Note that such an attack was studied in [16].

C. The Impact of Periodic LRDoS Attacks

As it is easy for an attacker to launch a periodic LRDoS attack that has a constant interval between attack pulses, we analyze the impact of such attack on the web server and leave the theoretical investigation of non-periodic LRDoS attacks to future work. In Section VI and the supplementary material, we evaluate different non-periodic LRDoS attacks and compare them with the periodic LRDoS attacks through experiments.

We quantify the impact of a periodic LRDoS attack from three aspects:

- Proposition 4 proves that the victim system's state oscillates along with the attack.
- Proposition 5 quantifies the range of the admission rate under an LRDoS attack.
- Proposition 6 and Proposition 7 characterize the effectiveness of an LRDoS attack.

Proposition 4: Under a periodic LRDoS attack, the web server's state converges with a periodic solution.

Proposition 5: If the web server is under a periodic LRDoS attack, the maximal and minimal values of the admission rate are shown in Eqn. (5) and (6). They increase with τ and decrease with $\lambda_a^{\mathcal{W}}$.

$$\alpha_{max} = \begin{cases} \left(\alpha(\eta^{\mathcal{W},1}) - \frac{\rho^*-B}{A\lambda_n^{\mathcal{W}}}\right)e^{-A\lambda_n^{\mathcal{W}}K(\tau-\eta^{\mathcal{W},1})} + \frac{\rho^*-B}{A\lambda_n^{\mathcal{W}}}, & \text{type I attack,} \\ \left(\alpha(\eta^{\mathcal{W},2}) - \frac{\rho^*-D}{C\lambda_n^{\mathcal{W}}}\right)e^{-C\lambda_n^{\mathcal{W}}K(\tau-\eta^{\mathcal{W},1}-\eta^{\mathcal{W},2})} + \frac{\rho^*-D}{C\lambda_n^{\mathcal{W}}}, & \text{type II attack,} \\ \alpha^c, & \text{type III attack,} \end{cases} \quad (5)$$

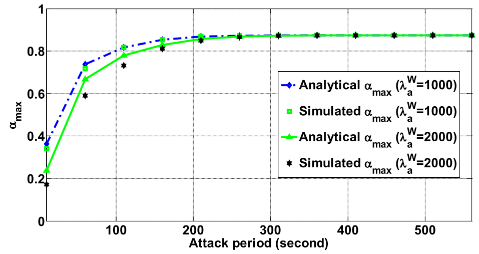


Fig. 5. α_{max} for different values of τ and $\lambda_a^{\mathcal{W}}$.

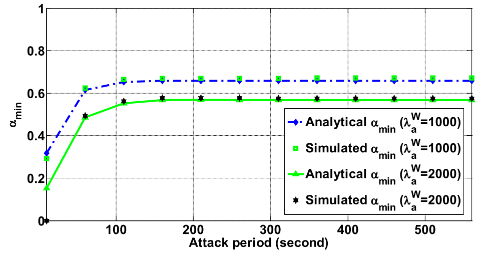


Fig. 6. α_{min} for different values of τ and $\lambda_a^{\mathcal{W}}$.

where α^c is the desired value when $\rho = \rho^*$, $\alpha(\eta^{\mathcal{W},1}) = K(\rho^* - 1)\eta^{\mathcal{W},1} + \alpha_0$ and $\alpha(\eta^{\mathcal{W},2}) = (\alpha(\eta^{\mathcal{W},1}) - \frac{\rho^*-B}{A\lambda_n^{\mathcal{W}}})e^{-A\lambda_n^{\mathcal{W}}K\eta^{\mathcal{W},2}} + \frac{\rho^*-B}{A\lambda_n^{\mathcal{W}}}$.

$$\alpha_{min} = K(\rho^* - 1)\eta^{\mathcal{W},1} + \alpha_{max}. \quad (6)$$

We verify α_{max} (i.e., Eqn. (5)) and α_{min} (i.e., Eqn. (6)) through simulation using the parameters from [16]. Fig. 5 and Fig. 6 show that the analytical results closely match with the simulation results. We can see that both α_{max} and α_{min} increase with τ. The reason is that for the type I and II attacks a larger τ will result in less damage. As shown in Fig.5, with the same τ, a larger $\lambda_a^{\mathcal{W}}$ will cause smaller α_{max} and α_{min} because it imposes severer damage to the server. For the type III attack, since $\alpha(t)$ can restore to the steady value, α_{max} equals to that value regardless $\lambda_a^{\mathcal{W}}$. As shown in Fig. 6, although α_{min} will also converge to the same value when τ increases, a larger $\lambda_a^{\mathcal{W}}$ will lead to a smaller α_{min} because it forces the server to further decrease the admission rate. Comparing Fig.5 and Fig.6, we can see that it takes a longer time for α_{max} to converge than α_{min}. It is because the recovery procedure consists of two long stages whereas $\alpha(t)$ quickly drops to α_{min} on the arrival of attack pulses.

We define two metrics to characterize the effectiveness of an LRDoS attack after the system converges:

- The percentage of normal requests dropped due to the attack. $\phi = \frac{\int_0^\tau (\alpha^c - \alpha(t))\lambda_n^{\mathcal{W}} dt}{\int_0^\tau \alpha^c \lambda_n^{\mathcal{W}} dt}$, where α^c is the admission rate when the system is in the steady state in the absence of LRDoS attacks.
- The number of dropped normal requests per attack request. $\psi = \frac{\int_0^\tau (\alpha^c - \alpha(t))\lambda_n^{\mathcal{W}} dt}{\int_0^\tau \lambda_a^{\mathcal{W}} \delta(t) dt}$.

Propositions 6 and 7 prove that ϕ is a decreasing function of the attack period τ while ψ is an increasing function of the attack period τ, respectively. That is, to drop more normal requests, the attacker should adopt a shorter period. In the extreme case, the attacker launches a flooding attack, which leads to large cost in terms of sending more requests and the high risk of being detected. We will investigate the optimal attack strategy considering both the effectiveness and the cost of LRDoS attack in future work.

Proposition 6: After the system converges, ϕ is a decreasing function of τ.

Proposition 7: After the system converges, ψ is an increasing function of τ.

D. A Switched System Model

The above analysis shows that the web server experiences several stages on the arrival of attack pulses. Therefore, we model the system under attack as a switched system. We use $\alpha(t)$ to represent the system state because it directly controls the incoming requests and the switched model for $\alpha(t)$ is:

$$\dot{\alpha}(t) = K(\rho^* - \rho(t)) \tag{7}$$

Section IV-B indicates that $\rho(t)$ has different trajectories and switching points in $t \in [T_k, T_{k+1})$. For the type I attack, $\eta_k^{\mathcal{W},1} < \tau_{k+1} < \eta_k^{\mathcal{W},1} + \eta_k^{\mathcal{W},2}$, we have

$$\rho(t) = \begin{cases} 1, & T_k \le t < T_k + \eta_k^{\mathcal{W},1}, \\ A\lambda\alpha(t) + B, & T_k + \eta_k^{\mathcal{W},1} \le t < T_{k+1}. \end{cases}$$

There are two series of switching points: T_k denotes time-dependent switching points caused by the arrival of the k^{th} attack pulse, and $s_j = T_k + \eta_k^{\mathcal{W},1}$ represents state-dependent switching points due to the event $\rho(t_j) < \rho^*$.

For the type II attack, $\eta_k^{\mathcal{W},1} + \eta_k^{\mathcal{W},2} \le \tau_{k+1} < \eta_k^{\mathcal{W},1} + \eta_k^{\mathcal{W},2} + \eta_k^{\mathcal{W},3}$, we have

$$\rho(t) = \begin{cases} 1, & T_k \le t < T_k + \eta_k^{\mathcal{W},1}. \\ A\lambda\alpha(t) + B, & T_k + \eta_k^{\mathcal{W},1} \le t < T_k + \eta_k^{\mathcal{W},1} + \eta_k^{\mathcal{W},2}, \\ C\lambda\alpha(t) + D, & T_k + \eta_k^{\mathcal{W},1} + \eta_k^{\mathcal{W},2} \le t < T_{k+1}. \end{cases}$$

There are three series of switching points: T_k, $s_j = T_k + \eta_k^{\mathcal{W},1}$, and $s_{j+1} = T_k + \eta_k^{\mathcal{W},1} + \eta_k^{\mathcal{W},2}$. The first one represents time-dependent switching points resulted from the arrival of the k^{th} attack pulse while the other two are state-dependent switching points caused by the event $\rho(s_j) < \rho^*$ and $n(s_{j+1}) = \upsilon$.

For the type III attacks, $\tau_{k+1} \ge \eta_k^{\mathcal{W},1} + \eta_k^{\mathcal{W},2} + \eta_k^{\mathcal{W},3}$, we have

$$\rho(t) = \begin{cases} 1, & T_k \le t < T_k + \eta_k^{\mathcal{W},1}, \\ A\lambda\alpha(t) + B, & T_k + \eta_k^{\mathcal{W},1} \le t < T_k + \eta_k^{\mathcal{W},1} + \eta_k^{\mathcal{W},2}, \\ C\lambda\alpha(t) + D, & T_k + \eta_k^{\mathcal{W},1} + \eta_k^{\mathcal{W},2} \le t < T_k + \eta_k^{\mathcal{W},1} \\ & \quad + \eta_k^{\mathcal{W},2} + \eta_k^{\mathcal{W},3}, \\ \rho^*, & T_k + \eta_k^{\mathcal{W},1} + \eta_k^{\mathcal{W},2} + \eta_k^{\mathcal{W},3} \le t < T_{k+1}. \end{cases}$$

There are four series of switching points: T_k, $s_j = T_k + \eta_k^{\mathcal{W},1}$, $s_{j+1} = T_k + \eta_k^{\mathcal{W},1} + \eta_k^{\mathcal{W},2}$, and $s_{j+2} = T_k + \eta_k^{\mathcal{W},1} + \eta_k^{\mathcal{W},2} + \eta_k^{\mathcal{W},3}$. The first one denotes time-dependent switching points due to the arrival of the k^{th} attack pulse while the other three represent state-dependent switching points caused by the event $\rho(s_j) < \rho^*$, $n(s_{j+1}) = \upsilon$, and $\rho(s_{j+1}) = \rho^*$, respectively.

E. Stability Analysis

In Section III we prove the existence of LRDoS attacks that can make the system under attack Lyapunov and Lagrange stable. Here, the stability analysis provides the following new results, including (1) the Lyapunov function (i.e., Eqn. (8)) for proving the system's stability; (2) the condition for an LRDoS attack to make the web server Lyapunov stable (Proposition 8); (3) the relationship between the bound of the system trajectory and the parameters of an LRDoS attack (Proposition 9);

We rewrite Eqn. (7) as follows: (1) $\dot{\alpha}(t) = f_1(\alpha) = K(\rho^* - 1)$; (2) $\dot{\alpha}(t) = f_2(\alpha) = K(\rho^* - A\lambda\alpha - B)$; (3) $\dot{\alpha}(t) = f_3(\alpha) = K(\rho^* - C\lambda\alpha - D)$; (4) $\dot{\alpha}(t) = f_4(\alpha) = 0$.

The roots of $f_p = 0$, $p \in \{1, 2, 3, 4\}$ are the equilibrium points, including $\alpha_{e1} = 0$, $\alpha_{e2} = \frac{\rho^* - B}{A\lambda}$, $\alpha_{e3} = \frac{\rho^* - D}{C\lambda}$ and $\alpha_{e4} = \alpha^c$. The system is Lyapunov stable if all the equilibrium points are Lyapunov stable. In the following analysis we ignore the case of f_1 because $\dot{\alpha}(t) = K(\rho^* - 1)$ implies that the admission rate continuously decreases until $\alpha(t) = 0$. In other words, the admission rate converges to 0.

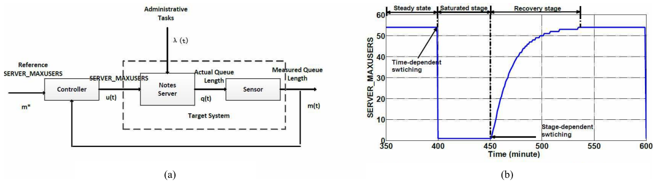
To make the origin be the equilibrium point such that $f_p(0) = 0$ for all $p \in \{2, 3, 4\}$ [44], we let $x(t) = \alpha(t) - \alpha_{ep}$ for $\sigma(t) = p$, where α_{ep}, $p = \{2\}$ for type I attack, $p = \{2, 3\}$ for type II attack, and $p \in \{2, 3, 4\}$ for type III attack, is the equilibrium point of the switched system (7).

Proposition 8 establishes the connection between the Lyapunov stability and the sequence of attack pulses by identifying conditions on the attack periods. That is each interval between two consecutive attack pulses should be larger than the saturation period (i.e., $T_{k+1} - T_k > \eta_k^{\mathcal{W},1}$ for all k). If $T_{k+1} - T_k \le \eta_k^{\mathcal{W},1}$, $\alpha(t)$ converges to zero.

Besides proving that the web server under attack is Lagrange stable, Proposition 9 further establishes the relationship between the bound of the system state and the parameters of an LRDoS attack by proving $\alpha(t)$ is bounded in the neighborhood of α_{ep}. This result shows that an attacker can tune an LRDoS attack for causing different degree of damage to the web server.

Proposition 8: Let the multiple Lyapunov function $V_p(\alpha - \alpha_{\mathrm{e}p})$ be

$$V_p(\alpha - \alpha_{\mathrm{e}p}) = (\alpha - \alpha_{\mathrm{e}p})^2. \tag{8}$$



(a)

(b)

Fig. 7. Attacking an IBM Notes server. (a) The IBM Notes server model. (b) System output in the presence of an LRDoS attack.

The switched system (7) is Lyapunov stable if $\tau_k > \eta_k^{\mathcal{W},1}$ for all k.

Proposition 9: Given the switched system (7) is Lyapunov stable, it is Lagrange stable, and

- If $\eta_k^{\mathcal{W},1} < \tau_{k+1} < \eta_k^{\mathcal{W},1} + \eta_k^{\mathcal{W},2}$, (i.e., Type I attack), $\alpha(t)$ is bounded by
$$\alpha(t) \leq (K(\rho^* - 1)\eta_k^{\mathcal{W},1} + \alpha_{e2} - \frac{\rho^* - B}{A\lambda_n^{\mathcal{W}}})e^{-A\lambda_n^{\mathcal{W}}K(\tau_k - \eta_k^{\mathcal{W},1})} + \frac{\rho^* - B}{A\lambda_n^{\mathcal{W}}}, k \in \mathbb{R}^n.$$
- If $\eta_k^{\mathcal{W},1} + \eta_k^{\mathcal{W},2} \leq \tau_{k+1} < \eta_k^{\mathcal{W},1} + \eta_k^{\mathcal{W},2} + \eta_k^{\mathcal{W},3}$, (i.e., Type II attack), $\alpha(t)$ is bounded by
$$\alpha(t) \leq (\alpha' - \frac{\rho^* - D}{C\lambda_n^{\mathcal{W}}})e^{-C\lambda_n^{\mathcal{W}}K(\tau_k - \eta^{\mathcal{W},1} - \eta^{\mathcal{W},2})} + \frac{\rho^* - D}{C\lambda_n^{\mathcal{W}}},$$
where $\alpha' = (K(\rho^* - 1)\eta^{\mathcal{W},1} + \alpha_{e3} - \frac{\rho^* - B}{A\lambda_n^{\mathcal{W}}})e^{-A\lambda_n^{\mathcal{W}}K\eta^{\mathcal{W},2}} + \frac{\rho^* - B}{A\lambda_n^{\mathcal{W}}}$
- If $\tau_{k+1} \geq \eta_k^{\mathcal{W},1} + \eta_k^{\mathcal{W},2} + \eta_k^{\mathcal{W},3}$, (i.e., Type III attack), the system converges to $\alpha_{e4} = \rho^*$.

V. Attacking an IBM Notes Server

IBM Notes has been widely used in enterprise networks since its debut in 1989. In this section, we examine a feedback-control based IBM Notes server proposed in [21], which manages the tradeoffs between its response time and throughput by controlling queue length. There are three major difference between the IBM Notes server model ($\mathcal{N}$) and the web server model ($\mathcal{W}$) analyzed in Section IV. First, they serve for different applications. Note that $\mathcal{N}$ is an email server. Second, they adopt different control models. $\mathcal{W}$ uses the PI controller while $\mathcal{N}$ employs the I controller [2]. Third, attack pulses enter the system through different operators. $\mathcal{W}$ adopts a multiplicative operator while $\mathcal{N}$ uses an add operator. Despite these differences, our theoretical analysis and experimental results illustrate that the LRDoS attacks can also cause severe damage to the IBM Notes server.

A. The IBM Notes Server Model

Fig. 7(a) shows the IBM Notes server model [21], which uses SERVER_MAXUSERS to regulate the number of users permitted to access the server. After connecting to the server, a user can send many remote procedure call (RPC) requests. The number of in-process RPC requests is referred to as queue length. After setting a desired queue length, the system adjusts SERVER_MAXUSERS to make the real queue length, which is measured through a sensor, converge to the desired value. Parekh et al. suggested that in practice SERVER_MAXUSERS should not be less than one in [21].

Let $q(t)$, $m(t)$, $u(t)$ and m^* represent the actual queue length, the measured queue length, SERVER_MAXUSERS, and the desired queue length, respectively. The server employs an integral controller to tune SERVER_MAXUSERS (i.e., $u(t)$) according to the difference between the desired queue length (i.e., m^*) and the measured queue length (i.e., $m(t)$). Let R_n and R_a denote the arrival rate of RPCs sent by a normal user and an attacker, respectively. We use R to denote the total arrival rate of RPCs. We consider the service rate $\mu^{\mathcal{N}}$ explicitly in the model following [49]. The continuous model for the IBM Notes server is as follows:

$$\begin{cases} \dot{q}(t) = Ru(t) - \mu^{\mathcal{N}}, \quad q(t) \geq 0 \\ a_s\dot{m}(t) = (a_s - 1)m(t) + (b_{s1} + b_{s2})q(t) - b_{s2}\dot{q}(t), \\ \hfill m(t) \geq 0 \\ \dot{u}(t) = K_i(m^* - m(t)), \quad u(t) \geq 1 \end{cases}$$

where a_s, b_{s1}, b_{s2} and K_i are constants.

B. Different Types of LRDoS Attacks

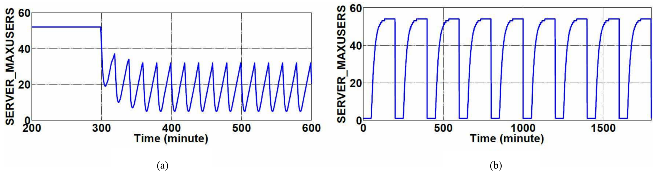
The goal of an LRDoS attack is to restrict SERVER_MAXUSERS so that new connections from legitimate users will be dropped. To achieve it, an attacker sends intermittent attack pulses composed of RPC requests through one or more established connections to inflate the queue length, which will force the system to decrease SERVER_MAXUSERS.

On the arrival of an attack pulse, the server evolves through two different stages before returning to the steady state: *saturation* and *recovery*, as shown in Fig. 3. We use $\eta^{\mathcal{N},1}$ and $\eta^{\mathcal{N},2}$ to denote the duration of each stage, respectively. The saturation stage begins right after the arrival of an attack pulse that increases $m(t)$ and drives $u(t)$ to $u_{min} = 1$ [21]. During the saturation stage, SERVER_MAXUSERS remains 1 and $m(t)$ continually decreases. After $m(t) < m^*$, the system enters the recovery stage, during which $m(t)$ first decreases due to small $u(t)$ and then increases towards m^*. The system state in these two stages are described in Eqn. (9) and (10), respectively.

$$\begin{cases} \dot{q}(t) = R - \mu^{\mathcal{N}}, \quad q(t) \geq 0 \\ a_s\dot{m}(t) = (a_s - 1)m(t) + (b_{s1} + b_{s2})q(t) - b_{s2}\dot{q}(t), \\ \hfill m(t) \geq 0 \\ \dot{u}(t) = 0, \quad u(t) = 1 \end{cases} \tag{9}$$

$$\begin{cases} \dot{q}(t) = Ru(t) - \mu^{\mathcal{N}}, \quad q(t) \geq 0 \\ a_s\dot{m}(t) = (a_s - 1)m(t) + (b_{s1} + b_{s2})q(t) - b_{s2}\dot{q}(t), \\ \hfill m(t) \geq 0 \\ \dot{u}(t) = K_i(m^* - m(t)), \quad u(t) > 1 \end{cases} \tag{10}$$

Saturation stage: Once an attack pulse arrives, the system enters the saturation stage that ends when $m(t) \leq m^*$. $\eta^{\mathcal{N},1}$



(a)

(b)

Fig. 8. Two types of LRDoS attacks according to the relationship between the attack period and the duration of two stages. (a) Type I attack. (b) Type II attack.

can be obtained by solving Eqn. (11).

$$\exp\left(\left(\frac{a_s + K_i b_{s1} - 1}{2a_s}t\right)\right. \times \left(\cosh\left(\frac{A_1^{0.5}}{a_s}t\right) + \frac{a_s A_3}{-(u^*)^{0.5}} \sinh\left(\frac{A_2^{0.5}}{a_s}t\right)\right) = 0, \quad (11)$$

where $\cosh(x)$ and $\sinh(x)$ are hyperbolic cosine and sine functions. A_1, A_2, and A_3 are a set of constants used in this section. Please find their expressions in the supplementary material because they are long and complicated.

Recovery stage: At the beginning of this stage, since $m(\eta^{\mathcal{N},1}) \leq m^*$, $u(t)$ increases. The queue length $m(t)$ keeps decreasing until $m(t) = \frac{b_{s1}+b_{s2}}{a_s-1}(u(t) - \dot{u}(t))$ and then increases. $u(t)$ increases until it equals to u^*, where $u^* = \frac{1-a_s}{b_{s1}+b_{s2}}m^*$ is the value of SERVER_MAXUSERS when the system is in steady state. $\eta^{\mathcal{N},2}$ is the solution of Eqn. (12).

$$\cosh(A_1\eta^{\mathcal{N},2}) - a_{s1}\sinh(A_1\eta^{\mathcal{N},2}) = 0 \quad (12)$$

The two stages indicates that $u(t)$ has two possible trajectories in $t \in [T_k, T_{k+1})$. In the saturation stage, $u(t) = 1$. In recovery stage, $u(t) = e^{A_0 t}\frac{\cosh(A_1 t) - a_{s1}\sinh(A_1 t)}{A_1 a_{s1} b_{s1}}A_3 + u^*$. According to the relationship between the attack period and the duration of two stages, we identify two types of attacks. Fig. 8 shows SERVER_MAXUSERS's trajectory under different types of LRDoS attacks.

Type I attack: It has $\tau_{k+1} < \eta_k^{\mathcal{N},1} + \eta_k^{\mathcal{N},2}$. Under such attack, SERVER_MAXUSERS's trajectory involves two stages: the saturation stage and the recovery stage, as illustrated in Fig.8(a). More precisely, a new attack pulse arrives when the system is still in the recovery stage.

When $\tau_{k+1} \leq \eta_k^{\mathcal{N},1}$, the next attack pulse arrives during the saturation stage. In this case, the system state oscillates between the saturation stage and the recovery stage for finite times $k < k_0$, where k_0 is a finite constant. The reasons behind such behavior are as follows. The $(k+1)^{th}$ attack pulse arrives when the system is still saturated due to the requests in the k^{th} attack pulse. Then, SERVER_MAXUSERS remains one, $q(t)$ keeps on decreasing as all new incoming users are rejected except one ($u(t) = 1$). Consequently, the measured queue $m(t)$ keeps on decreasing. If $\tau_{k+1} < \eta_k^{\mathcal{N},1}$ holds for all the attack pulses, $u(t)$ will be fixed at one. In the extreme case, it becomes the flooding-based attack. We exclude this case and redefine the type I attack as $\eta_k^{\mathcal{N},1} < \tau_{k+1} < \eta_k^{\mathcal{N},1} + \eta_k^{\mathcal{N},2}$.

Type II attack: It has $\tau_{k+1} \geq \eta_k^{\mathcal{N},1} + \eta_k^{\mathcal{N},2}$. The evolution of the system involves the saturation stage, the recovery stage, and the steady state, as illustrated in Fig. 8(b).

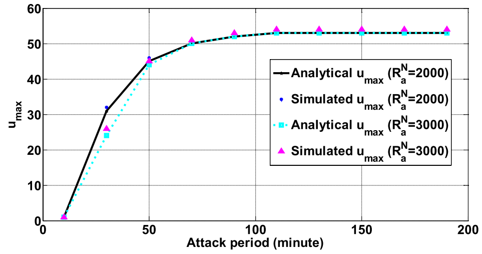


Fig. 9. Analytical and simulation results for u_{max} under LRDoS attacks with different parameters.

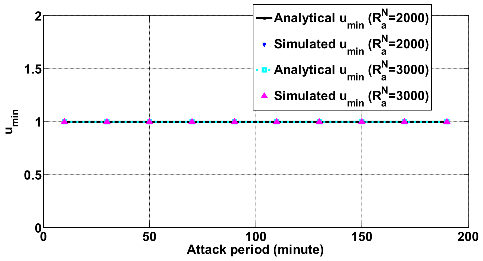


Fig. 10. Analytical and simulation results for u_{min} under LRDoS attacks with different parameters.

C. The Impact of Periodic LRDoS Attacks

Proposition 10 proves that the system's state oscillates along with the attack. Proposition 11 gives the maximal and minimal values of SERVER_MAXUSERS in the presence of a periodic LRDoS attack.

Proposition 10: Under a periodic LRDoS attack, the IBM Notes server's state converges with a periodic solution.

Proposition 11: If an IBM Notes server is under a periodic LRDoS attack, the minimal value of SERVER_MAXUSERS is one (i.e., $u_{min} = 1$) and its maximal value is:

$$u_{max} = \begin{cases} e^{A_0(\tau-\eta^{\mathcal{N},1})}\frac{\cosh(A_1(\tau-\eta^{\mathcal{N},1}))}{A_1 a_{s1} b_{s1}}A_3 \\ -e^{A_0(\tau-\eta^{\mathcal{N},1})}\frac{a_{s1}\sinh(A_1(\tau-\eta^{\mathcal{N},1}))}{A_1 a_{s1} b_{s1}}A_3 + u^*, \\ \qquad\qquad\qquad\qquad \text{Type I attack} \\ u^*, \qquad\qquad\qquad \text{Type II attack,} \end{cases}$$

We verify u_{max} and u_{min} through simulation using the parameters from [21]. Fig. 9 and Fig. 10 show the analytical results and the simulation results under LRDoS attacks with different parameters. We can see that the analytical results match well with the simulation results. When $\tau < 130$ minutes, the system is under the type I attack. As shown in Fig. 9, in this case, u_{max} increases with τ because a long attack interval means the system has more time to recover $u(t)$ toward the steady value (i.e., a large u_{max}). Moreover, a larger $R_a^{\mathcal{N}}$ results in a smaller u_{max} because it causes more damage to the system. By contrast, for the type II attack, different $R_a^{\mathcal{N}}$s result in the same u_{max}, because $u(t)$ restores to the steady value. As shown in Fig. 10, if $R_a^{\mathcal{N}}$ is large enough to drive the system to enter the saturation stage, $u_{min} = 1$ (i.e., the smallest value of $u(t)$ [21]). We demonstrate the result when $R_a^{\mathcal{N}}$ is small in the supplementary material and find that a small $R_a^{\mathcal{N}}$ cannot cause significant damage to the system. We define two metrics to quantify the effectiveness and the efficiency of an LRDoS attack, including $\phi^{\mathcal{N}} = \frac{\int_0^\tau (u^* - u(t))dt}{\int_0^\tau u^* dt}$ and $\psi^{\mathcal{N}} = \frac{\int_0^\tau (u^* - u(t))dt}{\int_0^\tau R_a^{\mathcal{N}} \delta(t)dt}$, where u^* is the desired value of $u(t)$.

D. A Switched System Model

The above analysis shows that the system experiences several states on the arrival of attack pulses. Therefore, we model the system under attack as a switched system. We use $u(t)$ to represent the system state because it directly controls the incoming connections and the switched model for $u(t)$ is:

$$\dot{u}(t) = f_p(u), \tag{13}$$

Section V-B shows that $u(t)$ has different trajectories and switching points in $t \in [T_k, T_{k+1})$. For the type I attack, $\eta_k^{\mathcal{N},1} \leq \tau_{k+1} < \eta_k^{\mathcal{N},1} + \eta_k^{\mathcal{N},2}$, we have

$$u(t) = \begin{cases} 1, & T_k \leq t < T_k + \eta_k^{\mathcal{N},1}, \\ e^{A_0 t} \frac{\cosh(A_1 t) - a_{s1} \sinh(A_1 t)}{A_1 a_{s1} b_{s1}} A_3 + u^*, & T_k + \eta_k^{\mathcal{N},1} \leq t < T_{k+1}. \end{cases}$$

There are two series of switching points: T_k denotes time-dependent switching points due to the arrival of the k^{th} attack pulse, and $s_j = T_k + \eta_k^{\mathcal{N},1}$ represents state-dependent switching points caused by the event $m(t_j) < m^*$.

For the type II attack, $\tau_{k+1} \geq \eta_k^{\mathcal{N},1} + \eta_k^{\mathcal{N},2}$, we have

$$u(t) = \begin{cases} 1, & T_k \leq t < T_k + \eta_k^{\mathcal{N},1}, \\ e^{A_0 t} \frac{\cosh(A_1 t) - a_{s1} \sinh(A_1 t)}{A_1 a_{s1} b_{s1}} A_3 & \\ +u^*, & T_k + \eta_k^{\mathcal{N},1} \leq t < T_{k+1}, \\ u^*, & T_k + \eta_k^{\mathcal{N},1} + \eta_k^{\mathcal{N},2} \leq t < T_{k+1}. \end{cases}$$

There are three series of switching points: T_k, $t_j = T_k + \eta_k^{\mathcal{N},1}$, and $t_{j+1} = T_k + \eta_k^{\mathcal{N},1} + \eta_k^{\mathcal{N},2}$. The first one denotes time-dependent switching points due to the arrival of the k^{th} attack pulse. Others represent state-dependent switching points caused by the event $m(s_j) < m^*$ and $m(t_{j+1}) = m^*$.

E. Stability Analysis

We obtain the following results through stability analysis, including (1) the Lyapunov function (i.e., Eqn. (14)) for proving the system's stability; (2) the condition for an LRDoS attack to make the IBM Notes server Lyapunov stable (Proposition 12); (3) the relationship between the bound of the system trajectory and the parameters of an LRDoS attack (Proposition 13);

We rewrite Eqn. (13) as follows: (1) $\dot{u}(t) = f_1(u) = 0$; (2) $\dot{u}(t) = f_2(u) = K_i(m^* - m(t))$; (3) $\dot{u}(t) = f_3(u) = 0$.

The roots of $f_p = 0$, $p \in \{1, 2, 3\}$ are the equilibriums points, including $u_{e1} = 1$, $u_{e2} = u^*$ and $u_{e3} = u^*$. We ignore the case of u_{e1} because $u(t) = 1$ is a constant.

To make the origin be the equilibrium point such that $f_p(0) = 0$ for all $p \in \{1, 2, 3\}$ [44], we let $x(t) = u(t) - u_{ep}$ for $\sigma(t) = p$, where u_{ep}, $p = \{2\}$ for the type I attack and $p = \{2, 3\}$ for the type II attack, is the equilibrium point of the switched system (13).

Proposition 12 proves that an LRDoS attack with intervals larger than the saturation period (i.e., $T_{k+1} - T_k > \eta_k^{\mathcal{N},1}$ for all k) can make the system Lyapunov stable. In other words, such LRDoS attacks can force the system to stay away from the steady state. Otherwise, the attack becomes the flooding-based DoS attacks and $u(t)$ converges to 1. After proving that the system under attack is Lagrange stable, Proposition 13 further establishes the relationship between the bound of the system state and the parameters of an LRDoS attack by proving $u(t)$ is bounded in the neighborhood of u_{ep}. Such result allows an attacker to tune an LRDoS attack for causing certain degree of damage to the system.

Proposition 12: Let the multiple Lyapunov function $V_p(u - u_{ep})$ be

$$V_p(u - u_{\mathrm{e}p}) = (u - u_{\mathrm{e}p})^2. \tag{14}$$

The switched system (13) is Lyapunov stable if $\tau_k > \eta_k^{\mathcal{N},1}$ for all k.

Proposition 13: Given the switched system (13) is Lyapunov stable, it is Lagrange stable, and

- If $\eta_k^{\mathcal{N},1} < \tau_{k+1} < \eta_k^{\mathcal{N},1} + \eta_k^{\mathcal{N},2}$, (i.e., Type I attack), $u(t)$ is bounded by

$$u(t) \leq e^{A_0(\tau - \eta_k^{\mathcal{N},1})} \frac{\cosh(A_1(\tau_k - \eta_k^{\mathcal{N},1})) - a_{s1} \sinh(A_1(\tau_k - \eta_k^{\mathcal{N},1}))}{A_1 a_{s1} b_{s1}} A_3 + u_{e2}, k \in \mathbb{R}^n.$$

- If $\tau_{k+1} \geq \eta_k^{\mathcal{N},1} + \eta_k^{\mathcal{N},2}$, (i.e., Type II attack), $u(t)$ converges to u^*.

VI. Experiments

We carry out extensive experiments to evaluate the LRDoS attacks on the two servers. For the web server, we use the parameters from [16]: $A = 0.00267$, $B = 0.2$, $C = 0.024$, $D = -1.4$, $v = 75$, $K = 0.01$, $\mu = 90$, and $\rho^* = 0.7$. For the IBM Notes server, we adopt the parameters from [21]: $a_s = 0.6371$, $b_{s1} = 0.1692$, $b_{s2} = -0.1057$ and $K_i = 0.1$. The major reason for using the parameters from the original paper is that they make the system stable in the absence of LRDoS attacks. It is worth noting that while practical feedback control systems are necessarily stable [18] how to turn the controllers' parameters to make different systems stable is still an active research topic [50], [51] and out of the scope of this paper. In these experiments, we vary the parameters of the LRDoS attacks and the input of legitimate users to the servers, and examine the corresponding impact. Section VI-A and Section VI-B present the Matlab simulation result of the web server and the IBM Notes server, respectively. Due to the page limit, we leave the testbed experiment results and many other simulation results in the supplementary material.

A. Simulation Result for the Web Server Model

In Section IV, we analyze the effect of the attack period τ on $\phi^{\mathcal{W}}$ (i.e., percentage of normal requests dropped) and $\psi^{\mathcal{W}}$ (i.e., number of normal requests dropped per attack request). We evaluate such effect by launching LRDoS attacks with $\lambda_a^{\mathcal{W}} = 1000, 2000$ requests per second and a wide range of attack periods. Note that the $\lambda_a^{\mathcal{W}}$s are high enough to force the web server to enter the saturation stage on the arrival of each

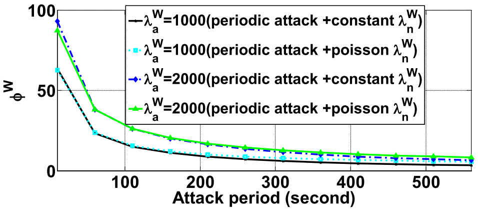


Fig. 11. The effectiveness of LRDoS attacks with fixed intervals between consecutive attack pulses: percentage of normal requests dropped.

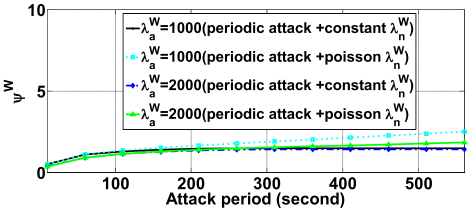


Fig. 12. The effectiveness of LRDoS attacks with fixed intervals between consecutive attack pulses: number of normal requests dropped per attack request.

attack pulse. We also examine small $\lambda_a^{\mathcal{W}}$s that cannot saturate the web server by individual attack pulse, and report the results in the supplementary material. In the experiments, the attacker sends either periodic pulses (in Fig. 11 and Fig. 12) or random pulses. For the latter, the interval between consecutive attack pulses follows exponential, normal and Pareto distributions. We only show the results of the exponential distribution in Fig. 13 and Fig. 14, and leave the others in the supplementary material. Moreover, we simulate both constant and Poisson arrival rate for normal requests.

Fig. 11 illustrates that $\phi^{\mathcal{W}}$ decreases with τ, because the web server has longer time to recover its admission rate and consequently takes in more normal requests. Hence, if an attacker wants to cause more legitimate requests to be dropped, she should use a smaller attack period. The extreme case is the flooding attack, whose period is zero. We also observe that $\phi^{\mathcal{W}}$ increases with $\lambda_a^{\mathcal{W}}$, because a larger $\lambda_a^{\mathcal{W}}$ causes severer damage. However, when we consider the attack cost (e.g., the number of attack request), a shorter interval or a larger $\lambda_a^{\mathcal{W}}$ may not be preferred. Fig. 12 shows that $\psi^{\mathcal{W}}$ increases with τ. That is, a larger attack period yields more damage per attack request. Moreover, $\psi^{\mathcal{W}}$ converges to a constant as τ increases, because in this situation the attacks belong to the type III attack and the web server can return to the steady state during the intervals between consecutive attack pluses. In other words, enlarging τ will not increase the number of dropped requests. Similarly, a larger $\lambda_a^{\mathcal{W}}$ may not be cost-effective as Fig. 12 shows that $\psi^{\mathcal{W}}$ decreases with $\lambda_a^{\mathcal{W}}$.

Fig. 11 and Fig. 12 illustrate the differences in the effectiveness of LRDoS attacks due to the different arrival processes of normal requests (i.e., constant-rate process versus Poisson process whose mean value is equal to the constant rate). We can observe from Fig. 11 and Fig. 12 that both $\phi^{\mathcal{W}}$ and $\psi^{\mathcal{W}}$ under the constant-rate process are less than those under the Poisson process. The reason may be that the bursts in the Poisson process were superposed on attack pulses and then caused more normal requests to be dropped, not to mention that the large bursts alone can also cause request losses in the absence of attack.

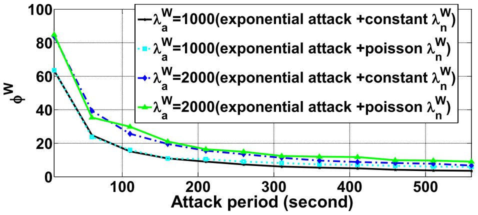


Fig. 13. The effectiveness of LRDoS attacks with exponentially distributed intervals between consecutive attack pulses: percentage of normal requests dropped.

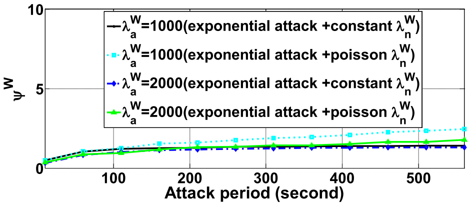


Fig. 14. The effectiveness of LRDoS attacks with exponentially distributed intervals between consecutive attack pulses: number of normal requests dropped per attack request.

Fig. 12 also shows that the difference for $\psi^{\mathcal{W}}$ under two arrival processes increases with τ. It may result from the different types of the attack. More precisely, when τ increases, the attack evolves from type I, type II to type III. Under the type III attacks, the total number of dropped normal requests caused by each attack pulse is fixed if the arrival rate of normal request remains constant. In contrast, the bursts in the Poisson arrival processes along with the attack pulses may force more normal requests to be discarded. Under the type I and II attacks (i.e., τ is small), the difference caused by different arrival processes is not obvious, because the attack pulses lead to the majority of dropped requests.

Fig. 13 and Fig. 14 illustrates similar results in terms of the relationship between $\phi^{\mathcal{W}}$ (or $\psi^{\mathcal{W}}$) with τ when the intervals between consecutive attack pulses follow the exponential distribution. However, comparing Fig. 13 and Fig. 11, we find that using the same $\lambda_a^{\mathcal{W}}$ the LRDoS attack with randomized interval causes less damage than the attack with fixed interval, especially for small τ. The reason may be that on one hand when the interval is longer than the fixed attack period the attack may allow the server to have more time to recover. On the other hand, when the interval is shorter than the fixed attack period, the current attack pulse may cause the server to drop more attack requests in the next attack pulse, thus moderating the attack damage. Moreover, similar to Fig. 12, Fig. 14 shows that the difference for $\psi^{\mathcal{W}}$ under two arrival processes of normal requests (i.e., constant-rate process versus Poisson process) becomes more obvious when τ increases.

B. Simulation Result for the IBM Notes Server Model

For evaluating the vulnerability of the IBM Notes server, we simulate normal users arriving at a rate of $\lambda_n = 200$ per minute and let the number of RPC requests sent by each normal user follow the Poisson distribution with a mean value of 5 RPCs, i.e. $\mathbb{E}[R_n^{\mathcal{N}}(t) = 5]$. In this

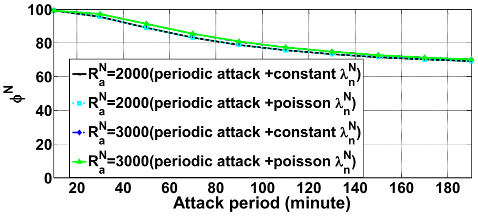


Fig. 15. The effectiveness of LRDoS attacks with fixed intervals between consecutive attack pulses: percentage of normal connections dropped.

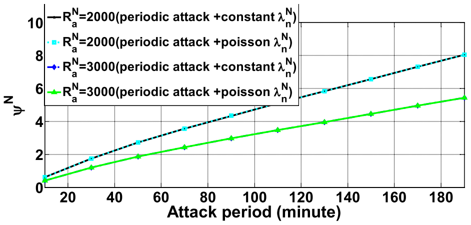


Fig. 16. The effectiveness of LRDoS attacks with fixed intervals between consecutive attack pulses: number of normal connections dropped per attack request.

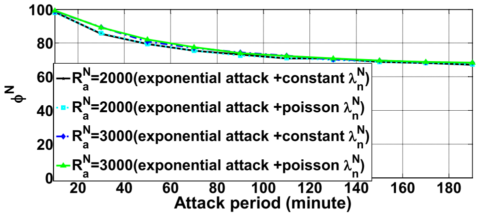


Fig. 17. The effectiveness of LRDoS attacks with exponentially distributed intervals between consecutive attack pulses: percentage of normal connections dropped.

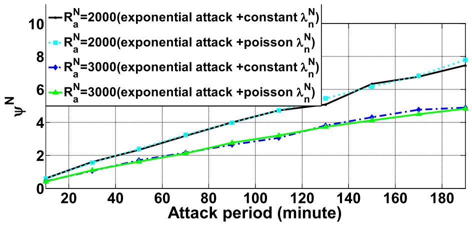


Fig. 18. The effectiveness of LRDoS attacks with exponentially distributed intervals between consecutive attack pulses: number of normal connections dropped per attack request.

experiment, there is only one attack connection because it is enough for an attacker to cause damage by dispatching high volume of RPCs in this connection. To make the attack more stealthy, an attacker can dispatch those RPCs through several connections. We vary the intervals between consecutive attack pulses and $R_a^{\mathcal{N}}$ (i.e., the number of RPC sent by the attacker) to evaluate the effectiveness of LRDoS attacks. More precisely, we let $R_a^{\mathcal{N}} = 2000, 3000$ requests per attack pulse and examine a wide range of intervals that are fixed or random. Note that these $R_a^{\mathcal{N}}$s are large enough to force the server to enter the saturation stage on the arrival of each attack pulse. The experimental results for small $R_a^{\mathcal{N}}$s are in the supplementary material.

Fig. 15 and Fig. 16 show the result for the LRDoS attacks with fixed attack period. Fig. 15 demonstrates that $\phi^{\mathcal{N}}$ increases with $R_a^{\mathcal{N}}$ and decreases with τ, meaning that to cause more severer damage an attacker had better shorten the attack period and adopt a larger $R_a^{\mathcal{N}}$. The reason is that a longer τ allows the server to increase SERVER_MAXUSERS to a larger value until it reaches the steady-state value as demonstrated in Fig. 8. In contrast, $\psi^{\mathcal{N}}$ decreases with $R_a^{\mathcal{N}}$ and increases with τ as shown in Fig. 16. Therefore, a small $R_a^{\mathcal{N}}$ and a large τ could increase the efficiency of each attack request.

Fig. 17 and Fig. 18 demonstrate that the different arrival processes of normal users did not lead to large difference in $\phi^{\mathcal{N}}$ (or $\psi^{\mathcal{N}}$). It is due to the intrinsic feature of the feedback controller in the IBM Notes server. More precisely, on one hand, when the number of incoming users exceeds SERVER_MAXUSERS, the Notes server will handle SERVER_MAXUSERS users and drop others no matter which distribution $\lambda_n^{\mathcal{N}}$ follows. Note that since the mean of the Poisson process is set to the same value as the constant-rate arrival process, they will generate almost the same amount of normal connections in a long period of time. On the other hand, when the number of incoming users is less than SERVER_MAXUSERS, $\lambda(t)$ users will be taken in. We call them active users. According to the system model in Section V-A, we know that it takes $a_s/(b_{s2}RK_i)$ users to reduce SERVER_MAXUSERS by one, because in stable stage one additional active user results in $b_{s2}R/a_s$ increment in the queue length and consequently $b_{s2}RK_i/a_s$ increment in SERVER_MAXUSERS. Therefore, if the bursts in the arrival process of normal requests are less than $a_s/(b_{s2}RK_i)$, there will not be significant differences in SERVER_MAXUSERS.

We also evaluate the impact of the LRDoS attacks with randomized intervals. Fig. 17 and Fig. 18 show the result for the attacks with intervals that follow the exponential distribution. Comparing Fig. 15 and Fig. 17 (or comparing Fig. 16 and Fig. 18), we find that the relationship between $\phi^{\mathcal{N}}$ (or $\psi^{\mathcal{N}}$) with $R_a^{\mathcal{N}}$ and τ is the same with periodic attacks. However, the impact in terms of $\phi^{\mathcal{N}}$ and $\psi^{\mathcal{N}}$ from the attack with randomized intervals is relatively less than those from the attack with fixed interval, especially in large τ. It may be due to the fact that the IBM Notes server just limits SERVER_MAXUSERS instead of dropping RPC requests sent by the attacker. In other words, one attack pulse will not cause the RPC requests in the consecutive attack pulse to be dropped. Hence, the slight difference may be just due to the randomized interval. The results with other settings such as different distributions can be found in the supplementary material.

VII. Discussion

The goal of this paper is to reveal the vulnerability of feedback-control based systems to the LRDoS attacks through theoretical analysis and then propose a new methodology to quantify the impact of the LRDoS attacks on such systems. Therefore, we assume that the feedback control model for the victim system is available, such as, the web server model in [16] and the IBM notes server model in [21]. However, some feedback-control based systems may not have constructed the model. For example, the system in [7] just measures the output (i.e., $y(t)$) of the *process* (i.e., $h(t)$ without giving the equations of $h(t)$. In order to apply our methodology to analyze such systems, users can first model $h(t)$ through system identification [52]. There are also systems that do not use feedback controller or employ other kinds of controllers, such as adaptive control, model predictive control, robust control, etc. Our methodology could not investigate the impact of LRDoS attacks on such systems and we will examine them in future work.

To simplify the theoretical analysis, we assume that the arrival rates of normal requests to the two servers are constants. Although this assumption may not be realistic, our analysis sheds light on the impact of the LRDoS attacks. Other arrival processes along with the attack may cause severer damage. The reason is that the bursts in the arrival process of normal users may be superposed on the attack pulses and consequently cause more damage to the server, not to mention that the large bursts alone may also affect the server. The experiment results in Section VI demonstrate it. In future work, we will enhance our model by considering more realistic arrival process models.

VIII. Conclusion

We investigate the vulnerability of feedback-control based Internet services to the LRDoS attacks. We first examine the impact of the LRDoS attacks on a general feedback control system and prove that LRDoS attacks can force the system's steady-state error to oscillate along with the attack. By modeling the system under attack as a switched system, we prove the existence of LRDoS attacks that can drive the system to a state other than the desired state. Both the oscillation of steady-state error and staying away from the desired state impair the system's performance. Then, we propose a novel methodology to analyze the impact of LRDoS attacks on specific feedback control systems. We obtain many new insights by applying the methodology to examine a web server model and an IBM Notes server model. In future work, we will investigate the tradeoff between the effectiveness and the cost of LRDoS attack, and design defense mechanism to mitigate the damage of LRDoS attack.

Acknowledgment

We thank Prof. T. Charles Clancy and the reviewers for their careful reading of the manuscript and for their very helpful suggestions.

References

[1] T. Abdelzaher, K. Shin, and N. Bhatti, "Performance guarantees for web server end-systems: A control-theoretical approach," *IEEE Trans. Parallel Distrib. Syst.*, vol. 13, no. 1, pp. 80–96, Jan. 2002.

[2] J. Hellerstein, Y. Diao, S. Parekh, and D. Tilbury, *Feedback Control of Computing Systems*. Hoboken, NJ, USA: Wiley, 2004.

[3] M. Welsh and D. Culler, "Adaptive overload control for busy internet servers," in *Proc. USENIX Symp. Internet Technol. Syst.*, 2003, pp. 1–4.

[4] Y. Lu, T. Abdelzaher, C. Lu, L. Sha, and X. Liu, "Feedback control with queueing-theoretic prediction for relative delay guarantees in web servers," in *Proc. 19th IEEE RTAS*, May 2003, pp. 208–217.

[5] H. Lim, S. Babu, J. Chase, and S. Parekh, "Automated control in cloud computing: Challenges and opportunities," in *Proc. 1st Workshop ACDC*, Jun. 2009, pp. 13–18.

[6] Y. Seung, T. Lam, L. Li, and T. Woo, "CloudFlex: Seamless scaling of enterprise applications into the cloud," in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 211–215.

[7] A. Sharifi, S. Srikantaiah, A. Mishra, M. Kandemir, and C. Das, "METE: Meeting end-to-end QoS in multicores through system-wide resource management," *ACM SIGMETRICS Perform. Eval. Rev.*, vol. 39, no. 1, pp. 13–24, Jun. 2011.

[8] T. Abdelzaher, J. Stankovic, C. Lu, R. Zhang, and Y. Lu, "Feedback performance control in software services," *IEEE Control Syst.*, vol. 23, no. 3, pp. 74–90, Jun. 2003.

[9] S. Park and M. Humphrey, "Predictable high-performance computing using feedback control and admission control," *IEEE Trans. Parallel Distrib. Syst.*, vol. 22, no. 3, pp. 396–411, Mar. 2011.

[10] Z. Wang, Y. Chen, D. Gmach, S. Singhal, B. Watson, W. Rivera, *et al.*, "AppRAISE: Application-level performance management in virtualized server environments," *IEEE Trans. Netw. Service Manag.*, vol. 6, no. 4, pp. 240–254, Dec. 2009.

[11] K. Kim and P. Kumar, "Cyber–physical systems: A perspective at the centennial," *Proc. IEEE*, vol. 100, no. 13, pp. 1287–1308, May 2012.

[12] M. Huebscher and J. McCann, "A survey of autonomic computing—Degrees, models, and applications," *ACM Comput. Surv.*, vol. 40, no. 3, pp. 1–7, Aug. 2008.

[13] A. Kuzmanovic and E. Knightly, "Low-rate TCP-targeted Denial-of-service attacks: The shrew vs. the mice and elephants," in *Proc. Conf. Appl., Technol., Archit., Protocols Comput. Commun.*, Aug. 2003, pp. 75–86.

[14] M. Guirguis, A. Bestavros, I. Matta, and Y. Zhang, "Exploiting the transients of adaptation for RoQ attacks on internet resources," in *Proc. 12th IEEE ICNP*, Oct. 2004, pp. 184–195.

[15] X. Luo, R. Chang, and E. Chan, "Performance analysis of TCP/AQM under Denial-of-service attacks," in *Proc. 13th IEEE Int. Symp. MASCOTS*, Sep. 2005, pp. 97–104.

[16] M. Guirguis, A. Bestavros, I. Matta, and Y. Zhang, "Reduction of quality (RoQ) attacks on internet end-systems," in *Proc. IEEE 24th Annu. Joint Conf. Comput. Commun. Soc.*, vol. 2. Mar. 2005, pp. 1362–1372.

[17] M. Guirguis, A. Bestavros, I. Matta, and Y. Zhang, "Adversarial exploits of end-systems adaptation dynamics," *J. Parallel Distrib. Comput.*, vol. 67, no. 3, pp. 318–335, 2007.

[18] K. Ogata, *Modern Control Engineering*, 5th ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 2009.

[19] J. Lunze and F. Lamnabhi-Lagarrigue, *Handbook of Hybrid Systems Control: Theory, Tools, Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2009.

[20] W. Haddad, V. Chellaboina, and S. Nersesov, *Impulsive and Hybrid Dynamical Systems: Stability, Dissipativity, and Control*. Princeton, NJ, USA: Princeton Univ. Press, 2006.

[21] S. Parekh, N. Gandhi, J. Hellerstein, D. Tilbury, T. Jayram, and J. Bigus, "Using control theory to achieve service level objectives in performance management," *Real-Time Syst.*, vol. 23, nos. 1–2, pp. 127–141, 2002.

[22] J. Mirkovic and P. Reiher, "A taxonomy of DDoS attack and DDoS defense mechanisms," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 34, no. 2, pp. 39–53, 2004.

[23] T. Peng, C. Leckie, and K. Ramamohanarao, "Survey of network-based defense mechanisms countering the DoS and DDoS problems," *ACM Comput. Surv.*, vol. 39, no. 1, pp. 1–3, 2007.

[24] G. Loukas and G. Oke, "Protection against denial of service attacks: A survey," *Comput. J.*, vol. 53, no. 7, pp. 1020–1037, 2010.

[25] X. Luo, E. Chan, and R. Chang, "Vanguard: A new detection scheme for a class of TCP-targeted Denial-of-service attacks," in *Proc. 10th IEEE/IFIP NOMS*, Apr. 2006, pp. 507–518.

[26] Y. Zhang, Z. Mao, and J. Wang, "Low-rate TCP-targeted DoS attack disrupts internet routing," in *Proc. NDSS*, 2007, pp. 1–15.

[27] M. Schuchard, A. Mohaisen, D. Kune, N. Hopper, Y. Kim, and E. Vasserman, "Losing control of the internet: Using the data plane to attack the control plane," in *Proc. NDSS*, 2011, pp. 1–15.

[28] M. Guirguis, A. Bestavros, I. Matta, and Y. Zhang, "Reduction of quality (RoQ) attacks on dynamic load balancers: Vulnerability assessment and design tradeoffs," in *Proc. 26th IEEE Int. Conf. Comput. Commun.*, May 2007, pp. 857–865.

[29] W. Chen, Y. Zhang, and Y. Wei, "The feasibility of launching reduction of quality(RoQ) attacks in 802.11 wireless networks," in *Proc. 14th IEEE ICPADS*, Dec. 2008, pp. 517–524.

[30] Y. He, Q. Cao, Y. Han, L. Wu, and T. Liu, "Reduction of quality (RoQ) attacks on structured peer-to-peer networks," in *Proc. IEEE IPDPS*, May 2009, pp. 1–9.

[31] G. Macia-Fernandez, J. Diaz-Verdejo, and P. Garcia-Teodoro, "Evaluation of a low-rate DoS attack against iterative servers," *Comput. Netw.*, vol. 51, no. 4, pp. 1013–1030, Mar. 2007.

[32] G. Macia-Fernandez, J. Diaz-Verdejo, and P. Garcia-Teodoro, "Mathematical model for low-rate DoS attacks against application servers," *IEEE Trans. Inf. Forensics Security*, vol. 4, no. 3, pp. 519–529, Sep. 2009.

[33] G. Maciá-Fernández, R. Rodriguez-Góomez, and J. E. Diaz-Verdejo, "Defense techniques for low-rate DoS attacks against application servers," *Comput. Netw.*, vol. 54, no. 15, pp. 2711–2727, May 2010.

[34] H. Sun, J. Lui, and D. Yau, "Defending against low-rate TCP attacks: Dynamic detection and protection," in *Proc. IEEE ICNP*, Oct. 2004, pp. 196–205.

[35] Y. Chen and K. Hwang, "Collaborative detection and filtering of shrew DDoS attacks using spectral analysis," *J. Parallel Distrib. Comput.*, vol. 66, no. 9, pp. 1137–1151, Jun. 2006.

[36] A. Shevtekar, K. Anantharam, and N. Ansari, "Low rate TCP Denial-of-service attack detection at edge routers," *IEEE Commun. Lett.*, vol. 9, no. 4, pp. 363–365, Apr. 2005.

[37] A. Shevtekar and N. Ansari, "A proactive test based differentiation technique to mitigate low rate DoS attacks," in *Proc. ICCCN*, Aug. 2007, pp. 639–644.

[38] G. Thatte, U. Mitra, and J. Heidemann, "Detection of low-rate attacks in computer networks," in *Proc. IEEE INFOCOM Workshops*, Apr. 2008, pp. 1–6.

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

[40] A. Shevtekar and N. Ansari, "A router-based technique to mitigate reduction of quality (RoQ) attacks," *Comput. Netw.*, vol. 52, no. 5, pp. 957–970, Apr. 2008.

[41] C. Zhang, Z. Cai, W. Chen, X. Luo, and J. Yin, "Flow level detection and filtering of low-rate DDoS," *Comput. Netw.*, vol. 56, no. 15, pp. 3417–3431, Oct. 2012.

[42] Y. Tang, "Countermeasures on application level low-rate denial of service attack," in *Proc. 14th Int. Conf. ICICS*, Oct. 2012, pp. 70–80.

[43] C. Lu, J. Stankovic, G. Tao, and S. Son, "Feedback control real-time scheduling: Framework, modeling, and algorithms," *J. Real-Time Syst.*, vol. 23, nos. 1–2, pp. 85–126, 2002.

[44] J. Slotine and W. Li, *Applied Nonlinear Control*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1991.

[45] D. Liberzon, *Switching in Systems and Control*. Boston, MA, USA: Birkhäuser, 2003.

[46] M. Vidyasagar, *Nonlinear Systems Analysis*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1978.

[47] C. Toumazou and N. Battersby, *Circuits and Systems Tutorials*. Hoboken, NJ, USA: Wiley, 1995.

[48] M. Branicky, "Multiple Lyapunov functions and other analysis tools for switched and hybrid systems," *IEEE Trans. Autom. Control*, vol. 43, no. 4, pp. 475–482, Apr. 1998.

[49] J. Hellerstein, Y. Diao, and S. Parekh, "A first-principles approach to constructing transfer functions for admission control in computing systems," in *Proc. 41st IEEE CDC*, vol. 2. Dec. 2002, pp. 2906–2912.

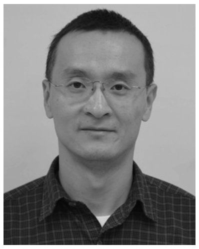
[50] A. McCormack and K. Godfrey, "Rule-based autotuning based on frequency domain identification," *IEEE Trans. Control Syst. Technol.*, vol. 6, no. 1, pp. 43–61, Jan. 1998.

[51] W. Levine, *The Control Handbook*, 2nd ed. Boca Raton, FL, USA: CRC Press, 2012.

[52] L. Ljung, *System Identification: Theory for the User*, 2nd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 1999.

Yajuan Tang received the Ph.D. degree in radio physics from Wuhan University in 2006. She is currently an Associate Professor with the Department of Electronic and Information Engineering, Shantou University. Her current research focuses on network security, network privacy, and malicious traffic analysis in networks using advanced signal processing techniques.

Xiapu Luo received the Ph.D. degree in computer science from Hong Kong Polytechnic University in 2007 and he was with the Georgia Institute of Technology as a Post-Doctoral Research Fellow. He is currently a Research Assistant Professor with the Department of Computing, Hong Kong Polytechnic University. He is a Researcher with the Shenzhen Research Institute, Hong Kong Polytechnic University. His current research focuses on network security and privacy, Internet measurement, and smartphone security.



Qing Hui received the Ph.D. degree in aerospace engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2008. In 2008, he joined the Department of Mechanical Engineering, Texas Tech University, Lubbock, TX, USA, where he is currently an Assistant Professor. His research interests include network robustness and vulnerability analysis, consensus, synchronization and control of network systems, network optimization, network interdependency and cascading failures, network threat detection, swarm optimization, hybrid systems, biomedical systems, and high performance scientific computing.

Rocky K. C. Chang received the Ph.D. degree in computer engineering from Rensselaer Polytechnic Institute. He joined the IBM Thomas J. Watson Research Center working on performance analysis and simulation tools. He joined the Department of Computing, Hong Kong Polytechnic University, where he is currently an Associate Professor. He is leading an Internet Infrastructure and Security Group, addressing problems in network security, network measurement, and network operations and management.