Optimal Security Patch Management Policies Maximizing System Availability

Toshikazu Uemura and Tadashi Dohi
Department of Information Engineering
Hiroshima University, Higashi-Hiroshima, Japan
Email: dohi@rel.hiroshima-u.ac.jp

Abstract— **In this paper we quantitatively evaluate dependability/security of a computer-based system subject to Denial of Service (DoS) attacks. More specifically, we develop two semi-Markov models for describing the stochastic behavior of systems with different security patch release strategies. The optimal security patch management policies are then formulated and analytically derived to maximize the steady-state system availability. We further perform the sensitivity analysis of model parameters through numerical experiments and refer to the effectiveness of our preventive patch management policies.**

Index Terms— **Security evaluation, availability, patch management policy, semi-Markov model, analytical approach.**

I. INTRODUCTION

The complexity, the heterogeneity and the openness of supporting infrastructures to untrusted internet users have given rise to an increasing number of vulnerabilities and malicious threats (viruses, worms, denial of service attacks, fishing attempts, *etc*). In general, there are three types of Internet attacks aimed against *confidentiality*, *integrity* and *availability* which are the most significant security attributes. Several attacks are aimed against confidentiality to obtain sensitive information such as commercial, industrial, political or even military secrets, as well as the personal data whose disclosure may endanger people privacy. Other attacks are aimed against the integrity of information; destruction or modification of sensitive data, spreading of false information, manipulation of published data, *etc*. Although Internet attacks are very common recently, the most frequent types of attacks are those aimed against availability, by *Denial of Service* (DoS) [20]. For malicious attackers, if the access right strengthens, the probability that the security intrusion happens will effectively decrease, but at the same time the utilization on accessibility will be rather lost. The classical security-related work has traditionally privileged, with a few exceptions, *intrusion avoidance* (vulnerability elimination, strong authentication, *etc*.) and *attack deterrence* (attack tracing, auditing, *etc*). However, such techniques have proved to be not sufficient to ensure the security of systems connected to the Internet.

More recently, *intrusion tolerance*, inspired from traditional dependable computing techniques commonly used for tolerating accidental faults in hardware and/or software systems, has received considerable attention to complement intrusion avoidance and has improved the security of systems connected to the Internet. So far, most efforts in security have been focused on specification, design and implementation issues. However, there is a lack of methods for objectively evaluating system behavior in the presence of malicious threats and quantifying the level of security achieved. Existing security evaluation techniques are currently qualitative, based more on the development process than on the developed products. Several implication techniques of intrusion tolerance at the architecture level have been developed for several real systems [8], [13], *e.g.*, distributed systems [16], middleware [14], database systems [10], [15], server systems [17]. The above approaches are based on the redundant design at the architecture level on secure software systems. In other words, these methods can be categorized by a design diversity technique in secure system design and need much cost for the development. On the other hand, the environment diversity technique by the temporal time redundancy is a low-cost security tolerance technique and seems to be quite effective in practice. In this paper we focus on the security patch management as a preventive maintenance which enables us to execute the temporal time redundancy, and investigate its effect on an intrusion tolerant system.

This paper is a continuation of the paper by the same authors (see [11]), where the DoS attacks are assumed. In the DoS attacks, the attackers detect vulnerabilities in a server application and make the network traffic increasing extremely by sending a large amount of illegal data. To protect the information assets from such malicious threats, the preventive action would be useful for tolerating the security failure. The typical but somewhat simple example of such a action may be a simple *security patch release*. If the vendors can know the vulnerable parts in a server application in advance, they can release the security patch before the malicious attackers detect them. In fact, the full vendors or the computer emergency response team/coordination center (CERT/CC) in the development side are always monitoring the system vulnerabilities reported by benign users or themselves, even after releasing the applications. In this paper, we develop two continuous-time semi-Markov chains (CTSMCs) for describing the stochastic behavior of an intrusion tolerant

Manuscript received February 28, 2009; revised October 7, 2009; accepted December 1, 2009.

doi:10.4304/jcm.5.1.71-80

system with different security patch timings. More specifically, we consider two cases where the vulnerability of an application software is detectable or not by the vulnerability identifiers. The former was discussed in the reference [11], the latter is treated in this paper in addition to the previous result. In the reference [9] we developed the expected cost modeling for only one semi-Markov model among two models. On the other hand, we are going to analyze the availability models with different patch management strategies in this paper.

The rest part of this paper is organized as follows: In Section II, we summarize the related work. In Section III and Section IV, we describe two CTSMCs which are referred to as Model 1 and Model 2. In Model 1, it is assumed that a vulnerable state of a system is detectable by vulnerability identifiers in the above scenario. On the other hand, in Model 2, it is assumed that the vendor can not always know the detection timing of vulnerabilities by malicious attackers. Based on the analytic technique of CTSMs, we derive the steady-state probabilities and the steady-state system availability for each model. Further, in both cases the optimal security patch management policies are analytically derived to maximize the steady-state system availability. Numerical examples are illustrated to show the sensitivity of model parameters in Section V. Finally, the paper is concluded in Section VI with some remarks.

II. RELATED WORK

The quantitative evaluation of information security based on modeling is, actually, effective to validate the effectiveness of computer-based systems with intrusion tolerance. Littlewood et al. [18] found the analogy between the information security theory and the traditional reliability theory in assessing the quantitative security of operational software systems, and proposed some quantitative security measures. Jonsson and Olovsson, [19] discussed a quantitative method to study the attacker's behavior with the empirical data observed in experiments. Ortalo et al. [4] applied the privilege graph and the continuous-time Markov chain (CTMC) to evaluate system vulnerabilities, and derived the mean effort to security failure. Singh et al. [6] and Stevens et al. [7] considered probabilistic models to verify the intrusion tolerant systems against several attack patterns, and explained theoretically the detection mechanism of system vulnerability. Madan et al. [1], [2] introduced an architecture with intrusion tolerance, called SITAR (Scalable Intrusion Tolerant Architecture) was considered, and described its stochastic behavior by a CTSMC. Uemura and Dohi [10] and Wang and Liu [15] discussed the availability and integrity optimization for an intrusion tolerant database system. The VoIP (Voice over IP) network system (see [5]) was modeled by CTMCs from the viewpoint of security design. Recently, Uemura *et al.* [12] extended Madan et al.'s results [1], [2] by introducing a control parameter. In this way, several stochastic models have been developed with the aim of quantitative evaluation of information security and system dependability.

III. MODEL 1

A. Notation and Definition

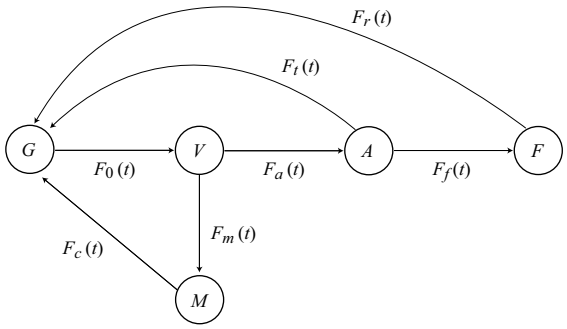

Figure 1. Semi-Markov transition diagram of Model 1.

Suppose that a server system starts operating at time $t = 0$ with *Normal State*; G. If attackers or hackers detect the vulnerability of a server application, the state makes a transition to *Vulnerable State*; V, where the transition time from G to V has a continuous cumulative distribution function (c.d.f.) $F_0(t)$ with mean μ_0 (> 0). Once a malicious attack begins, the system state changes to *Attack State*; A and the server operation stops for the corrective maintenance, where the transition time from V to A is given by a random variable having a continuous c.d.f. $F_a(t)$ and mean μ_a (> 0). In this phase, if the minor corrective maintenance in a failure probable state is performed such as data recovery, the system can be recovered from the failure probable state to the normal one, and can be considered to be as good as new. The transition time from State A to State G is given by the commonly distributed random variable with a c.d.f. $F_t(t)$ and mean μ_t (> 0). However, the system state may go to *System Failure State*; F before completing the minor corrective maintenance, where the transition time from A to F obeys a c.d.f. $F_f(t)$ with mean μ_f (> 0). Since this state is regarded as the system down state, the major recovery operation such as data initialization or system restart has to be carried out. The completion time to recover the server system from the system failure state is given by a non-negative continuous random variable with a c.d.f. $F_r(t)$ and mean μ_r (> 0).

On the other hand, if the vulnerable state V is detectable by vulnerability identifiers like a benign user, it may be effective to trigger the preventive patch management before the vulnerabilities are detected by malicious attackers. As an extreme scenario on preventive security patch management, suppose that a benign user discovers an application vulnerability faster than attackers, and discloses its information to the full vendor or the CERT/CC as well as his or her personal community. Then the security patch management is an important issue for the vendor. When the development period of patch is relatively shorter, is the quick release of the patch really beneficial? If the vulnerable state is seldom detected, it

would be better to release the security patch from the vendor as soon as possible. However, if the similar vulnerable states may come repeatedly, the frequent release of patches may lead to a large overhead in system operations. Define *Preventive Maintenance State*; M. If the preventive security patch management is triggered before the system becomes vulnerable, the system operation is stopped and the state goes to M from V. Without any loss of generality, we define the transition time from V to A having the following c.d.f.:

$$F_m(t) = \begin{cases} 1 & (t \geq t_0) \\ 0 & (t < t_0). \end{cases} \quad (1)$$

This means that the preventive release of a security patch is made at every t_0 time unit after the vulnerability is detected (this assumption can be relaxed in the latter discussion). Once the preventive patch management starts, it completes after the random time interval with a c.d.f. $F_c(t)$ and mean μ_c, so that the server system can be recovered similar to the state just before the vulnerability is detected. In the scenario on patch management, the time t_0 indicates a trigger to release the patch. The same cycle repeats again and again over an infinite time horizon. Since the underlying stochastic process is a CTSMC, we can apply the standard technique to study it. Figure 1 depicts the transition diagram of Model 1.

B. Behavioral Analysis

Suppose that the system state at time $t = 0$ is G with probability one. We define the transition probability from G to $j \in \{G, V, A, F, M\}$ at an arbitrary time t (> 0) and its Laplace-Stieltjes transform (LST) by $P_{Gj}(t)$ and $p_{Gj} = \int_0^\infty \exp\{-st\} dP_{Gj}(t)$, respectively. It is assumed that the underlying CTSMC is ergodic, *i.e.*, there exist the steady-state probabilities $\lim_{t\to\infty} P_{Gj}(t) = P_j$ $(j \in \{G, V, A, F, M\})$. Define the one-step transition probability of Model 1 and its LST by $Q_{ij}(t)$, $i, j \in \{G, V, A, F, M\}, i \neq j$ and $q_{ij}(s) = \int_0^\infty \exp\{-st\} dQ_{ij}(t)$, respectively. Then it is evident to obtain

$$\begin{aligned}
q_{GV}(s) &= \int_0^\infty \exp\{-st\} dF_0(t), & (2)\\
q_{VM}(s) &= \int_0^\infty \exp\{-st\} \overline{F}_a(t) dF_m(t), & (3)\\
q_{VA}(s) &= \int_0^\infty \exp\{-st\} \overline{F}_m(t) dF_a(t), & (4)\\
q_{AG}(s) &= \int_0^\infty \exp\{-st\} \overline{F}_f(t) dF_t(t), & (5)\\
q_{AF}(s) &= \int_0^\infty \exp\{-st\} \overline{F}_t(t) dF_f(t), & (6)\\
q_{FG}(s) &= \int_0^\infty \exp\{-st\} dF_r(t), & (7)\\
q_{MG}(s) &= \int_0^\infty \exp\{-st\} dF_c(t), & (8)
\end{aligned}$$

where in general $\overline{\psi}(\cdot) = 1 - \psi(\cdot)$.

Next we define the recurrent time distribution from State G to State G again by $H_{GG}(t)$. Then the LST of recurrent time distribution is given by

$$\begin{aligned}
h_{GG}(s) &= \int_0^\infty \exp\{-st\} dH_{GG}(t) \\
&= q_{GV}(s) q_{VA}(s) q_{AG}(s) \\
&\quad + q_{GV}(s) q_{VA}(s) q_{AF}(s) q_{FG}(s) \\
&\quad + q_{GV}(s) q_{VM}(s) q_{MG}(s). \quad (9)
\end{aligned}$$

From the result above, the LSTs of the transition probabilities, $p_{Gj}(s) = \int_0^\infty \exp\{-st\} dP_{Gj}(t)$, are given by

$$\begin{aligned}
p_{GG}(s) &= \overline{q}_{GV}(s)/\overline{h}_{GG}(s), & (10)\\
p_{GV}(s) &= q_{GV}(s)\Big(\overline{q}_{VA}(s) - q_{VM}(s)\Big)/\overline{h}_{GG}(s), & (11)\\
p_{GA}(s) &= q_{GV}(s) q_{VA}(s)\Big(\overline{q}_{AG}(s) - q_{AF}(s)\Big) \\
&\quad /\overline{h}_{GG}(s), & (12)\\
p_{GF}(s) &= q_{GV}(s) q_{VA}(s) q_{AF}(s) \overline{q}_{FG}(s)/\overline{h}_{GG}(s), & (13)\\
p_{GM}(s) &= q_{GV}(s) q_{VM}(s) \overline{q}_{MG}(s)/\overline{h}_{GG}(s). & (14)
\end{aligned}$$

It is not so easy to take the inversion of the above LSTs in Eqs.(10)–(14) analytically. Instead, by taking the limitation, we can derive the steady-state solutions $P_j = \lim_{t\to\infty} p_{G,j}(t), j \in \{G, V, A, F, M\}$ without proof:

Theorem 1.

$$\begin{aligned}
P_G &= \frac{\mu_0}{T_1(t_0)}, & (15)\\
P_V &= \frac{\int_0^{t_0} \overline{F}_a(t) dt}{T_1(t_0)}, & (16)\\
P_A &= \frac{\alpha F_a(t_0)}{T_1(t_0)}, & (17)\\
P_F &= \frac{\beta F_a(t_0)}{T_1(t_0)}, & (18)\\
P_M &= \frac{\mu_c \overline{F}_a(t_0)}{T_1(t_0)}, & (19)
\end{aligned}$$

where

$$\begin{aligned}
T_1(t_0) &= \mu_0 + \int_0^{t_0} \overline{F}_a(t) dt + \alpha F_a(t_0) \\
&\quad + \beta F_a(t_0) + \mu_c \overline{F}_a(t_0), \quad (20)\\
\alpha &= \int_0^\infty t \overline{F}_t(t) dF_f(t) \\
&\quad + \int_0^\infty t \overline{F}_f(t) dF_t(t), \quad (21)\\
\beta &= \mu_r \int_0^\infty \overline{F}_t(t) dF_f(t), \quad (22)
\end{aligned}$$

so that α and β in Eqs.(21) and (22) imply the mean transition time from State A to the subsequent state and the mean transition time from State A to State G through State F, respectively.

C. Optimal Security Patch Management Policy

The steady-state system availability for Model 1, $AV_{11}(t_0)$, is formulated by

$$\begin{aligned} AV_{11}(t_0) &= \lim_{t\to\infty} \frac{\mathrm{E}[\text{total UP time during } (0,t]]}{t} \\ &= P_G + P_V = U_{11}(t_0)/T_1(t_0), \end{aligned} \tag{23}$$

where

$$U_{11}(t_0) = \mu_0 + \int_0^{t_0} \overline{F}_a(t)dt. \tag{24}$$

This indicates the probability that the system subject to DoS attacks is operative in the steady state. On the other hand, it would be possible to consider the case where the patch is ready in advance for preventive maintenance even when the vulnerabilities are detected by malicious attackers. Then, it may be possible to operate the system without stopping the process in State A. For such a case, the steady-state system availability can be re-formulated as

$$AV_{12}(t_0) = P_G + P_V + P_A = U_{12}(t_0)/T_1(t_0), \tag{25}$$

where

$$U_{12}(t_0) = \mu_0 + \int_0^{t_0} \overline{F}_a(t)dt + \alpha F_a(t_0). \tag{26}$$

To distinguish two availability criteria, $AV_{11}(t_0)$ and $AV_{12}(t_0)$, we call respective models Model 1-1 and Model 1-2.

We make the following parametric assumptions:

(A-1) $\mu_c < \beta$.

Assumption **(A-1)** implies that the mean time to recover the system after a system failure is always greater than the mean time required by the preventive patch release. This is needed to motivate theoretically the optimal preventive patch management policy considered here. In Model 1-2, we further need the following technical assumption:

(A-2) $\alpha\mu_c < (\beta - \mu_c)\mu_0$.

Unfortunately, it is not easy to interpret the above assumption in terms of the cost component. In other words, the assumption (A-2) is technically required to guarantee the sufficiency of the optimal preventive patch management policy for Model 1-2.

Then, we can characterize the optimal preventive patch management policies maximizing the steady-state system availability in Model 1 (Model 1-1 and Model 1-2) as follows.

Theorem 2.

In Model 1-1, (1) Suppose that the c.d.f. $F_a(t)$ is strictly IFR (Increasing Failure Rate) under the assumption **(A-1)**, *i.e.*, the failure rate $r_a(t) = (dF_a(t)/dt)/\overline{F}_a(t)$ is strictly increasing in t. Define the non-linear function:

$$\begin{aligned} q_{11}(t_0) &= T_1(t_0) - \{1 + (\alpha + \beta - \mu_c) \\ &\quad \times r_a(t_0)\}U_{11}(t_0). \end{aligned} \tag{27}$$

(i) If $q_{11}(0) > 0$ and $q_{11}(\infty) < 0$, then there exists a finite and unique optimal preventive patch management policy t_0^* $(0 < t_0^* < \infty)$ satisfying $q_{11}(t_0^*) = 0$. The maximum steady-state system availability is then given by

$$AV_{11}(t_0^*) = \frac{1}{\{1 + (\alpha + \beta - \mu_c)\}r_a(t_0^*)}. \tag{28}$$

(ii) If $q_{11}(0) \le 0$, then $t_0^* = 0$, *i.e.*, it is optimal to release the preventive patch just after the vulnerability is detected. Then the maximum steady-state system availability is given by

$$AV_{11}(0) = \frac{\mu_0}{\mu_0 + \mu_c}. \tag{29}$$

(iii) If $q_{11}(\infty) \ge 0$, then $t_0^* \to \infty$, *i.e.*, it is optimal not to release the preventive patch even after the vulnerability is detected. Then the maximum steady-state system availability is given by

$$AV_{11}(\infty) = \frac{\mu_0 + \mu_a}{\mu_0 + \mu_a + \alpha + \beta}. \tag{30}$$

(2) Suppose that the c.d.f. $F_a(t)$ is DFR (Decreasing Failure Rate) under the assumption **(A-1)**, *i.e.*, the failure rate $r_a(t)$ is decreasing in t. If $AV_{11}(0) > AV_{11}(\infty)$, then $t_0^* = 0$ otherwise $t_0^* \to \infty$.

Proof: Differentiating the function $AV_{11}(t_0)$ with respect to t_0 and setting it equal to zero imply $q_{11}(t_0) = 0$. Further differentiation of $q_{11}(t_0)$ yields

$$\frac{dq_{11}(t_0)}{dt_0} = -\frac{dr_a(t_0)}{dt}(\alpha + \beta - \mu_c)U_{11}(t_0). \tag{31}$$

If $F_a(t)$ is strict IFR, from the assumption **(A-1)**, it is obvious that the right-hand-side of Eq.(31) takes a negative value for an arbitrary t_0 and that the function $q_{11}(t_0)$ is a decreasing function of t_0. From this monotone property, the steady-state system availability $AV_{11}(t_0)$ is a strictly quasi-convex function of t_0, so that if $q_{c1}(0) > 0$ and $q_{c1}(\infty) < 0$, then there exists a finite and unique optimal solution t_0^* $(0 < t_0^* < \infty)$ which satisfies $q_{11}(t_0^*) = 0$. In the cases of $q_{11}(0) \le 0$ and $q_{11}(\infty) \ge 0$, the steady-state system availability $AV_{11}(t_0)$ becomes increasing and decreasing in t_0, and the optimal solution is given by $t_0^* = 0$ and $t_0^* \to \infty$, respectively. If $F_a(t)$ is DFR, the system availability $AV_{11}(t_0)$ is a quasi-concave function of t_0, and the result is trivial.

Theorem 3.

In Model 1-2, (1) Suppose that the c.d.f. $F_a(t)$ is strictly IFR under the assumptions **(A-1)** and **(A-2)**. Define the non-linear function:

$$\begin{aligned} q_{12}(t_0) &= \{1 + \alpha r_a(t_0)\}T_1(t_0) \\ &\quad -\{1 + (\alpha + \beta - \mu_c)r_a(t_0)\}U_{11}(t_0). \end{aligned} \tag{32}$$

(i) If $q_{12}(0) > 0$ and $q_{12}(\infty) < 0$, then there exists a finite and unique optimal preventive

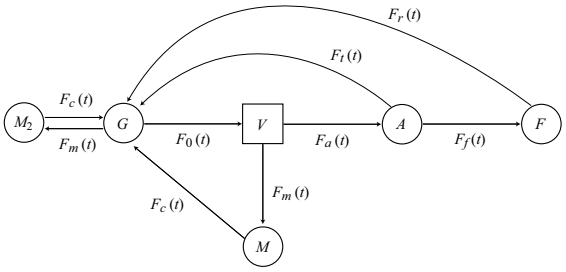

Figure 2. MRGP transition diagram of Model 2.

patch management policy t_0^* $(0 < t_0^* < \infty)$ satisfying $q_{12}(t_0^*) = 0$. The maximum steady-state system availability is then given by

$$AV_{12}(t_0^*) = \frac{1 + \alpha r_a(t_0^*)}{\{1 + (\alpha + \beta - \mu_c)\} r_a(t_0^*)}. \quad (33)$$

(ii) If $q_{12}(0) \leq 0$, then $t_0^* = 0$, and the maximum steady-state system availability is given by

$$AV_{12}(0) = \frac{\mu_0}{\mu_0 + \mu_c}. \quad (34)$$

(iii) If $q_{12}(\infty) \geq 0$, then $t_0^* \to \infty$, and the maximum steady-state system availability is given by

$$AV_{12}(\infty) = \frac{\mu_0 + \mu_a + \alpha}{\mu_0 + \mu_a + \alpha + \beta}. \quad (35)$$

(2) Suppose that the c.d.f. $F_a(t)$ is DFR under the assumptions **(A-1)** and **(A-2)**. If $AV_{12}(0) > AV_{12}(\infty)$, then $t_0^* = 0$ otherwise $t_0^* \to \infty$.

Proof: Differentiating the function $AV_{12}(t_0)$ with respect to t_0 and setting it equal to zero imply $q_{12}(t_0) = 0$. Further differentiation of $q_{12}(t_0)$ yields

$$\frac{q_{12}(t_0)}{dt_0} = \frac{dr_a(t_0)}{dt}\{\alpha T_1(t_0) - (\alpha + \beta - \mu_c) U_{12}(t_0)\}. \quad (36)$$

If $F_a(t)$ is strict IFR, from the assumptions **(A-1)** and **(A-2)**, the term of $\alpha T_1(t_0) - (\alpha + \beta - \mu_c) U_{12}(t_0)$ in the right-hand side of Eq.(36) is negative, so that

$$\alpha T_1(t_0) - (\alpha + \beta - \mu_c) U_{12}(t_0) = \alpha\mu_c - (\beta - \mu_c)\Big\{\mu_0 + \int_0^{t_0} \overline{F}_a(t)dt\Big\} < 0 \quad (37)$$

and $q_{12}(t_0)$ is a decreasing function of t_0. Hence, the steady-state system availability $AV_{12}(t_0)$ is quasi-convex and quasi-concave in t_0 for strictly IFR and DFR cases, respectively. The proof is completed.

IV. MODEL 2

A. Definition

In Model 1 it was assumed that the vulnerable state V is detectable by the vulnerability identifiers and that the development of the patch can be started after the time t_0 measured from the vulnerable state V elapsed. However, this assumption seems to be somewhat strong and can not be always validated, because the vendor can not always know the detection timing of vulnerabilities by malicious attackers. To resolve this problem, we consider another stochastic model referred as Model 2 in Fig.2, where another preventive maintenance state, M_2, is defined. That is, the preventive maintenance is triggered at the periodic time interval measured from State G. In Fig.2, the circles and the square denote regeneration points and a non-regeneration point, respectively, so that the underlying stochastic process is reduced to a Markov regenerative process (MRGP) which belongs to the wider class than the CTSMCs. However, as well known, the MRGP can be translated to the usual CTSMC by changing the definition of the underlying states. Figure 3 illustrates the translated CTSMC transition diagram of the MRGP in Fig.2, where we define two new states:

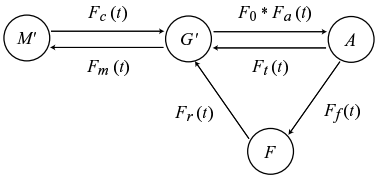

Figure 3. Translated CTSMC transition diagram of Model 2.

$G^{'}$; *Normal State*
$M^{'}$; *Aggregated Preventive Maintenance State*.

Define the Stieltjes convolution operator by '$*$', *i.e.*,

$$G(t) = F_0 * F_a(t) = \int_0^t F_0(t - x) dF_a(x). \quad (38)$$

Similar to the previous discussion in Model 1, suppose that the underlying MRGP is ergodeic, *i.e.*, if there exist the steady-state probabilities $\lim_{t\to\infty} P_{Gj}(t) = P_j$ $(j \in \{G', V, A, F, M'\})$. Then, it is straightforward to obtain the following result.

Theorem 4.

$$P_{G'} = \frac{\int_0^{t_0} \overline{G}(t)dt}{T_2(t_0)}, \quad (39)$$

$$P_A = \frac{\alpha G(t_0)}{T_2(t_0)}, \quad (40)$$

$$P_F = \frac{\beta G(t_0)}{T_2(t_0)}, \quad (41)$$

$$P_{M'} = \frac{\mu_c \overline{G}(t_0)}{T_2(t_0)}, \quad (42)$$

where

$$T_2(t_0) = \int_0^{t_0} \overline{G}(t)dt + \alpha G(t_0) + \beta G(t_0) + \mu_c \overline{G}(t_0). \quad (43)$$

B. Optimal Security Patch Management Policy

In Model 2, the steady-state system availability $AV_{21}(t_0)$ is formulated as

$$AV_{21}(t_0) = \lim_{t\to\infty} \frac{\mathrm{E}[\text{total UP time during } (0, t]]}{t}$$

$$= \quad P_G = U_{21}(t_0)/T_2(t_0), \tag{44}$$

where

$$U_{21}(t_0) = \int_0^{t_0} \overline{G}(t)dt. \tag{45}$$

On the other hand, when one focuses on the case where the patch is ready for preventive maintenance, in a fashion similar to Model 1, an alternative definition of the system availability is possible to consider. Similar to $AV_{12}(t_0)$ for Model 1-2, we define the steady-state system availability as

$$AV_{22}(t_0) = P_{G'} + P_A = U_{22}(t_0)/T_2(t_0), \tag{46}$$

where

$$U_{22}(t_0) = \int_0^{t_0} \overline{G}(t)dt + \alpha G(t_0). \tag{47}$$

We call the above models Model 2-1 and Model 2-2 corresponding to $AV_{21}(t_0)$ and $AV_{22}(t_0)$, respectively.

We give the following results to characterize the optimal preventive patch management policies for Model 2 (Model 2-1 and Model 2-2).

Theorem 5.

In Model 2-1, (1) Suppose that the c.d.f. $G(t)$ is strictly IFR under the assumption **(A-1)**. Define the non-linear function:

$$\begin{aligned} q_{21}(t_0) \quad = \quad & T_2(t_0) - \{1 + (\alpha + \beta - \mu_c) \\ & \times r_{0a}(t_0)\} U_{21}(t_0), \end{aligned} \tag{48}$$

where $r_{0a}(t) = (dG(t)/dt)/\overline{G}(t)$ is the failure rate.

(i) If $q_{c2}(\infty) < 0$, then there exists a finite and unique optimal preventive patch management policy t_0^* $(0 < t_0^* < \infty)$ satisfying $q_{21}(t_0^*) = 0$. The maximum steady-state system availability is then given by

$$AV_{21}(t_0^*) \quad = \quad \frac{1}{1 + (\alpha + \beta - \mu_c) r_{0a}(t_0^*)}. \tag{49}$$

(ii) If $q_{c2}(\infty) \geq 0$, then $t_0^* \to \infty$ and the maximum steady-state system availability is given by

$$AV_{21}(\infty) = \frac{\mu_0 + \mu_a}{\mu_0 + \mu_a + \alpha + \beta}. \tag{50}$$

(2) Suppose that the c.d.f. $G(t)$ is DFR under the assumption **(A-1)**. Then, we have $t_0^* \to \infty$.

Theorem 6.

In Model 2-2, (1) Suppose that the c.d.f. $G(t)$ is strictly IFR under the assumptions **(A-1)** and **(A-2)**. Define the non-linear function:

$$\begin{aligned} q_{22}(t_0) = & \{1 + \alpha r_{0a}(t_0)\} T_2(t_0) \\ & - \{1 + (\alpha + \beta - \mu_c) r_{0a}(t_0)\} U_{21}(t_0) \end{aligned} \tag{51}$$

and

$$A^* = \frac{\alpha}{\alpha + \beta}. \tag{52}$$

(i) If $AV_{22}(\infty) < A^*$, then the optimal preventive patch release timing t_0^* $(0 < t_0^* < \infty)$ with $q_{22}(t_0^*) = 0$ is strictly larger than $\tilde{t}$ satisfying $AV_{22}(\tilde{t}) = A^*$. Then, the maximum steady-state system availability is given by

$$AV_{22}(t_0^*) \quad = \quad \frac{1 + \alpha r_{0a}(t_0^*)}{1 + (\alpha + \beta - \mu_c) r_{0a}(t_0^*)}. \tag{53}$$

(ii) If $AV_{22}(\infty) \geq A^*$, then the optimal preventive patch management policy is given by $t_0^* \to \infty$ and its associated system availability becomes

$$AV_{22}(\infty) = \frac{\mu_0 + \mu_a + \alpha}{\mu_0 + \mu_a + \alpha + \beta}. \tag{54}$$

(2) Suppose that the c.d.f. $G(t)$ is DFR under the assumptions **(A-1)** and **(A-2)**.

(i) If $AV_{22}(\infty) < A^*$, then there exists a finite and unique optimal preventive patch management policy t_0^* $(0 < t_0^* < \infty)$ satisfying $AV_{22}(t_0^*) = A^*$ and $q_{22}(t_0^*) = 0$.

(ii) If $AV_{22}(\infty) \geq A^*$, then the optimal preventive patch management policy is given by $t_0^* \to \infty$ and its associated system availability is given by Eq.(54).

In Section III and Section IV, we derived the optimal preventive patch management policies for respective availability models with aperiodic and periodic patch release schedules, respectively. In the following section, we calculate numerically the optimal preventive patch management policies and their associated steady-state system availability, and compare them quantitatively. Also, we perform the sensitivity analysis of model parameters and investigate the effect of preventive maintenance policy in the intrusion tolerant system.

V. NUMERICAL ILLUSTRATIONS

Suppose that the c.d.f. $F_a(t)$ is given by the gamma distribution with shape parameter k (> 0) and scale parameter λ (> 0):

$$F_a(t) = 1 - \int_t^{\infty} \frac{x^{k-1} \lambda^k \exp\{-\lambda x\}}{\Gamma(k)} dx, \tag{55}$$

where $\Gamma(\cdot)$ is the standard gamma function. The other transition probabilities are given by the exponential distributions, where the model parameters are assumed as $\mu_0 = 168$, $\mu_f = 4$, $\mu_c = 3$, $\mu_t = 5$, $c_r = 2500$, $c_m = 500$ and $c_t = 750$. The main reason why the gamma distribution is assumed is due to its flexibility, so that the gamma distribution possesses both IFR and DFR properties for $k > 1$ and $k < 1$, respectively. When $k = 1$ it reduces the exponential distribution with memoryless property. Of our concern here is the investigation of sensitivity of shape parameter k and scale parameter λ on the optimal preventive patch management policies. Table I and Table II present the dependency of the parameters (k, λ) on the optimal patch management

TABLE I.
DEPENDENCE OF PARAMETERS (k, λ) ON STEADY-STATE SYSTEM AVAILABILITY (MODEL 1-1 V.S. MODEL 2-1).

		Model1			Model2		
(k,d)	$t_0 \to \infty$	t_0^*	$AV_{11}(t_0^*)$	inc	t_0^*	$AV_{21}(t_0^*)$	inc
(2,5)	0.9440	0.0598	0.9825	4.0720	120.2480	0.9510	0.7442
(2,10)	0.9468	0.2419	0.9825	3.7631	127.0030	0.9535	0.7064
(2,15)	0.9494	0.5504	0.9825	3.4861	133.7590	0.9558	0.6723
(3,5)	0.9455	0.8339	0.9825	3.9184	111.8050	0.9566	1.1816
(3,10)	0.9494	2.4427	0.9826	3.5001	120.9700	0.9598	1.0957
(3,15)	0.9528	4.6081	0.9828	3.1455	130.1340	0.9625	1.0214
(4,5)	0.9468	2.4225	0.9826	3.7806	111.7610	0.9608	1.4747
(4,10)	0.9517	6.3853	0.9829	3.2793	123.6500	0.9644	1.3379
(4,15)	0.9558	11.3065	0.9832	2.8761	135.5400	0.9675	1.2244
(5,5)	0.9481	4.5560	0.9828	3.6558	114.3100	0.9640	1.6735
(5,10)	0.9538	11.3713	0.9833	3.0906	129.1170	0.9680	1.4877
(5,15)	0.9584	19.4659	0.9838	2.6569	143.9240	0.9712	1.3391

TABLE II.
DEPENDENCE OF PARAMETERS (k, λ) ON STEADY-STATE SYSTEM AVAILABILITY (MODEL 1-2 V.S. MODEL 2-2).

		Model1			Model2		
(k,d)	$t_0 \to \infty$	t_0^*	$AV_{12}(t_0^*)$	inc	t_0^*	$AV_{22}(t_0^*)$	inc
(2,5)	0.9558	0.0858	0.9825	2.7889	158.4090	0.9583	0.2641
(2,10)	0.9580	0.3487	0.9825	2.5514	167.1740	0.9604	0.2515
(2,15)	0.9600	0.7975	0.9825	2.3388	175.9390	0.9623	0.2401
(3,5)	0.9569	1.0137	0.9825	2.6728	136.7910	0.9620	0.5234
(3,10)	0.9600	2.9944	0.9827	2.3549	147.8930	0.9647	0.4871
(3,15)	0.9627	5.6869	0.9828	2.0872	158.9940	0.9671	0.4555
(4,5)	0.9580	2.7934	0.9827	2.5709	131.5520	0.9649	0.7206
(4,10)	0.9619	7.4252	0.9830	2.1946	145.4450	0.9682	0.6563
(4,15)	0.9651	13.2298	0.9834	1.8951	159.3370	0.9709	0.6025
(5,5)	0.9591	5.1160	0.9828	2.4801	131.3010	0.9674	0.8658
(5,10)	0.9635	12.8703	0.9834	2.0604	148.2110	0.9710	0.7728
(5,15)	0.9671	22.1537	0.9840	1.7428	165.1210	0.9739	0.6979

TABLE III.
SYSTEM AVAILABILITY IN MODEL 1 FOR VARYING μ_t.

			Model1-1		Model1-2	
μ_t	α	β	t_0^*	$AV_{11}(t_0^*)$	t_0^*	$AV_{12}(t_0^*)$
1	0.8000	10.0000	0.3386	0.9825	0.3793	0.9825
2	1.3333	16.6667	0.1738	0.9825	0.1914	0.9825
3	1.7143	21.4286	0.1290	0.9825	0.1414	0.9825
4	2.0000	25.0000	0.1081	0.9825	0.1182	0.9825
5	2.2222	27.7778	0.0960	0.9825	0.1048	0.9825
6	2.4000	30.0000	0.0881	0.9825	0.0961	0.9825
7	2.5455	31.8182	0.0825	0.9825	0.0900	0.9825
8	2.6667	33.3333	0.0784	0.9825	0.0855	0.9825
9	2.7692	34.6154	0.0752	0.9825	0.0820	0.9825
10	2.8571	35.7143	0.0727	0.9825	0.0792	0.9825

TABLE IV.
SYSTEM AVAILABILITY IN MODEL 1 FOR VARYING μ_f.

			Model1-1		Model1-2	
μ_r	α	β	t_0^*	$AV_{11}(t_0^*)$	t_0^*	$AV_{12}(t_0^*)$
1		0.7143	7.0058	0.9828	∞	0.9963
2		1.4286	2.3810	0.9826	∞	0.9927
3		2.1429	1.4334	0.9825	∞	0.9891
4		2.8571	1.0253	0.9825	∞	0.9855
5	2.8571	3.5714	0.7980	0.9825	8.1050	0.9828
6		4.2857	0.6532	0.9825	2.4981	0.9826
7		5.0000	0.5529	0.9825	1.4752	0.9825
8		5.7143	0.4793	0.9825	1.0465	0.9825
9		6.4286	0.4230	0.9825	0.8108	0.9825
10		7.1429	0.3785	0.9825	0.6618	0.9825

policies and their associated system availability for Model 1 and Model 2, respectively. From these results, it is seen that the preventive maintenance is effective to improve the system availability. More precisely, in Model 1-1 and Model 2-1, the system availability was improved up to $2.66\% \sim 4.07\%$ and $0.67\% \sim 1.67\%$ by releasing the security patch preventively. On the other hand, the improvement for Model 1-2 was $1.74\% \sim 2.79\%$, but Model 2-2 showed $0.24\% \sim 0.87\%$ improvement. When the availability of the server system is kept to the high level such as five nines, these improvement may be significant, so that the preventive patch management would be effective to design the intrusion tolerant systems.

In Tables III-VI we compare the steady-state system

TABLE V.
SYSTEM AVAILABILITY IN MODEL 2 FOR VARYING μ_t.

μ_t	α	β	Model2-1		Model2-2	
			t_0^*	$AV_{21}(t_0^*)$	t_0^*	$AV_{22}(t_0^*)$
1	0.8000	10.0000	126.9110	0.9538	137.3420	0.9560
2	1.3333	16.6667	85.2727	0.9393	90.1283	0.9417
3	1.7143	21.4286	72.8766	0.9318	76.5945	0.9343
4	2.0000	25.0000	66.6848	0.9270	69.9089	0.9296
5	2.2222	27.7778	62.9235	0.9236	65.8700	0.9263
6	2.4000	30.0000	60.3816	0.9212	63.1497	0.9239
7	2.5455	31.8182	58.5434	0.9192	61.1868	0.9220
8	2.6667	33.3333	57.1496	0.9177	59.7011	0.9205
9	2.7692	34.6154	56.0554	0.9164	58.5360	0.9192
10	2.8571	35.7143	55.1728	0.9154	57.5971	0.9182

TABLE VI.
SYSTEM AVAILABILITY IN MODEL 2 FOR VARYING μ_f.

μ_r	α	β	Model2-1		Model2-2	
			t_0^*	$AV_{21}(t_0^*)$	t_0^*	$AV_{22}(t_0^*)$
1		0.7143	∞	0.9817	∞	0.9963
2		1.4286	∞	0.9782	∞	0.9927
3		2.1429	698.5970	0.9746	∞	0.9891
4		2.8571	352.9440	0.9712	∞	0.9855
5	2.8571	3.5714	257.6690	0.9680	∞	0.9820
6		4.2857	210.7800	0.9651	∞	0.9785
7		5.0000	182.1210	0.9625	798.2100	0.9750
8		5.7143	162.4750	0.9602	371.3680	0.9716
9		6.4286	148.0100	0.9580	266.1900	0.9684
10		7.1429	136.8270	0.9559	216.0210	0.9655

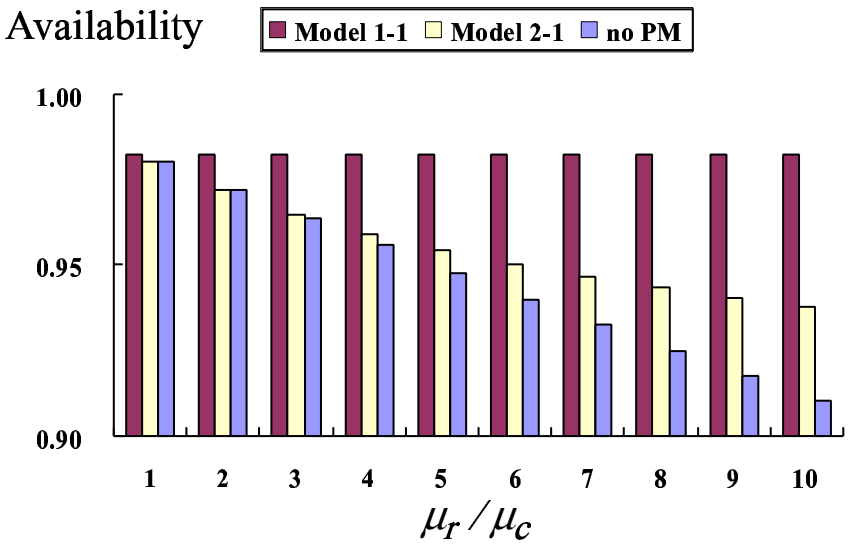

Figure 4. Comparison with steady-state system availability (Model 1-1 v.s. Model 2-1).

availability for varying μ_t and μ_r, which are related to the system recovery from preventive patch management and system failure. It can be shown that the mean recovery time μ_t strongly depends on the steady-state system availability and influences to the system availability as well as the system maintainability. Figures 4 and 5 illustrate the dependency of the parameter ratio μ_r/μ_c on the maximum steady-state system availability, where 'no PM' denotes the case of $t_0 \to \infty$. From these observations, it can be shown that Model 1 is always better than Model 2 in terms of the availability maximization. This seems to be a natural conclusion, because the transition to vulnerable state can be observed in Model 1. In this case, one sees that $3.02\% \sim 7.91\%$ improvement can be expected in system availability, if it is feasible. Hence, when the transition information of vulnerable state for the server system is available, it is possible to improve the system availability up to the maximum 7.91% in this setting. In other words, since Model 1 is an extreme case, the corresponding system availability can be regarded as an upper bound in the cases with feasible patch management policy.

VI. CONCLUSIONS

In this paper we have quantitatively evaluated an intrusion tolerant system with temporal diversity subject

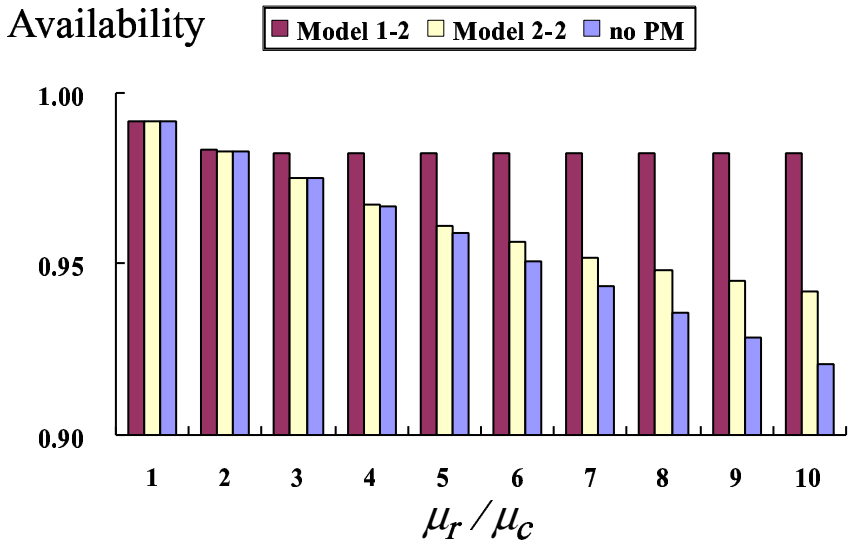

Figure 5. Comparison with steady-state system availability (Model 1-2 v.s. Model 2-2).

to DoS attacks in terms of the maximization of system availability, and developed two semi-Markov availability models to describe their stochastic behavior with different preventive patch management policies. We have derived the optimal preventive maintenance policies analytically to maximize the steady-state system availability. Further we have carried out the sensitivity analysis of some model parameters in numerical experiments, and have shown that the preventive patch release might be effective to improve the system availability in some cases.

In the future research, we are planning to estimate the model parameters in the operational phase of a real system, and to show the quantitative effectiveness based on the resulting CTSMC models. For instance, it is worth mentioning that the intrusion tolerant database system by [15] is quite similar but somewhat simpler than our models. Since the above authors did not take account of the preventive patch management for designing the intrusion tolerant database system, our approach based on the temporal diversity may be useful to improve the system availability.

ACKNOWLEDGMENT

This paper is based on "A Security Patch Management Model" by T. Uemura and T. Dohi, which appeared in the Proceedings of The 2009 IASTED International Conference on Software Engineering, pp. 114–119, ACTA Press (2009). This work was supported in part by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (C), Grant No. 19510148 (2007-2008).

REFERENCES

[1] B. B. Madan, K. Goseva-Popstojanova, K. Vaidyanathan, and K. S. Trivedi, "Modeling and quantification of security attributes of software systems,", *Proceedings of 32nd Annual IEEE/IFIP International Conference on Dependable Systems and Networks* (DSN 2002), pp. 505–514, IEEE CS Press (2002).

[2] B. B. Madan, K. Goseva-Popstojanova, K. Vaidyanathan, and K. S. Trivedi, "A method for modeling and quantifying the security attributes of intrusion tolerant systems,", *Performance Evaluation*, **56** (1/4), pp. 167–186 (2004).

[3] D. M. Nikol, W. H. Sanders and K. S. Trivedi, "Model-based evaluation: from dependability to security," *IEEE Transactions on Dependability and Secure Computing*, **1** (1), pp. 48–65 (2004).

[4] R. Ortalo, Y. Deswarte and M. Kaaniche, "Experimenting with quantitative evaluation tools for monitoring operational security," *IEEE Transactions on Software Engineering*, **25** (5), pp. 633–650 (1999).

[5] H. Pant, A. R. McGee, U. Chandrashekhar and S. H. Richman, "Optimal availability and security for IMS-based VoIP networks," *Bell Labs Technical Journal*, **11** (3), pp. 211–223 (2006).

[6] S. Singh, M. Cukier and W. H. Sanders, "Probabilistic validation of an intrusion tolerant replication system," *Proceedings of 33rd Annual IEEE/IFIP International Conference on Dependable Systems and Networks* (DSN 2003), pp. 615–624, IEEE CS Press (2003).

[7] F. Stevens, T. Courtney, S. Singh, A. Agbaria, J. F. Meyer, W. H. Sanders and P. Pal, "Model-based validation of an intrusion-tolerant information system," *Proceedings of 23rd IEEE Reliable Distributed Systems Symposium* (SRDS 2004), pp. 184–194, IEEE CS Press (2004).

[8] R. Stroud, I. Welch, J. Warne and P. Ryan, "A qualitative analysis of the intrusion-tolerant capabilities of the MAFTIA architecture," *Proceedings of 34th Annual IEEE/IFIP International Conference on Dependable Systems and Networks* (DSN 2004), pp. 453–461, IEEE CS Press (2004).

[9] T. Uemura and T. Dohi, "Quantitative evaluation of intrusion tolerant systems subject to DoS attacks via semi-Markov cost models," *Emerging Directions in Embedded and Ubiquitous Computing: International Conference EUC 2007 Workshops* (M. K. Denko, C.-S. Shih, K.-C. Li, S.-L. Tsao, Q.-A. Zeng, S.-H. Park, Y.-B. Ko, S.-H. Hung and J.-H. Park, eds.), LNCS **4809**, pp. 31–42, Springer-Verlag (2007).

[10] T. Uemura and T. Dohi, "Optimizing security measures in an intrusion tolerant database system," *ISAS 2008* (T. Nanya, F. Maruyama, A. Pataricza and M. Malek, eds.),

LNCS **5017**, pp. 26–42, Springer-Verlag (2008).
[11] T. Uemura and T. Dohi, "A security patch management model," *Proceedings of The 2009 IASTED International Conference on Software Engineering*, pp. 114–119, ACTA Press (2009).
[12] T. Uemura, T. Dohi and N. Kaio, "Availability analysis of an intrusion tolerant distributed server system with preventive maintenance," *IEEE Transactions on Reliability*, (in press).
[13] P. E. Verissimo, N. F. Neves and M. Correia, "Intrusion-tolerant architectures: concepts and design," *Architecting Dependable Systems (R. Lemos, C. Gacek and A. Romanovsky, eds.)*, LNCS **2677**, pp. 3–36, Springer-Verlag (2003).
[14] P. E. Verissimo, N. F. Neves, C. Cachin, J. Poritz, D. Powell,Y. Deswarte, R. Stroud and I. Welch, "Intrusion-tolerant middleware," *IEEE Security and Privacy*, **4** (4), pp. 54–62 (2006).
[15] H. Wang and P. Liu, "Modeling and evaluationg the survivability of an intrusion tolerant database system," *ESORICS 2006 (D. Gollmann, J. Meier and A. Sabelfeld, eds.)*, LNCS **4189**, pp. 207–224, Springer-Verlag (2006).
[16] Y. Deswarte, L. Blain and J. C. Fabre, "Intrusion tolerance in distributed computing systems," *Proceedings of 1991 IEEE Symposium on Research in Security and Privacy*, pp. 110–121, IEEE CS Press (1991).
[17] V. Guputa, V. Lam, H. V. Ramasamy, W. H. Sanders and S. Singh, "Dependability and performance evaluation of intrusion-tolerant server architectures," *LADC 2003*, LNCS **2847**, pp. 81–101, Springer-Verlag (2003).
[18] B. Littlewood, S. Brocklehurst, N. Fenton, P. Mellor, S. Page, D. Wright, J. Doboson, J. McDermid and D. Gollmann, "Towards operational measures of computer security," *Journal of Computer Security*, **2** (2/3), pp. 211–229, (1993).
[19] E. Jonsson and T. Olovsson, "A quantitative model of the security intrusion process based on attacker behavior," *IEEE Transactions on Software Engineering*, **23** (4), pp. 235–245, (1997).
[20] L. Garber, "Denial-of-service attacks rip the Internet," *IEEE Computer*, **33** (4), pp. 12–17, (2000).

Toshikazu Uemura received both the B.Sc. (Engineering) and M.Sc. (Engineering) from Hiroshima University, Japan in 2007 and 2009, respectively. Currently, he is working as a technical stuff in Mutsubishi Electric Corporation, Japan. His research interests in Hiroshima University were software security modeling and analysis.

Tadashi Dohi received the B.Sc. (Engineering), M.Sc. (Engineering), and Ph.D. (Engineering) from Hiroshima University, Japan, in 1989, 1991, and 1995, respectively. In 1992, he joined the Department of Industrial and Systems Engineering, Hiroshima University, Japan, as an Assistant Professor. Now he is a Full Professor at the Department of Information Engineering, Graduate School of Engineering, Hiroshima University, Japan, since 2002. In 1992, and 2000, he was a Visiting Research Scholar at University of British Columbia, Canada, and Duke University, USA, respectively, on leave of absence from Hiroshima University. His research areas include software reliability engineering, dependable computing, and performance evaluation. He is a Regular Member of ORSJ, JSIAM, IEICE, ISCIE, and IEEE.