Security Investment Analysis with Information Threats: A Stochastic Approach

WON SEOK YANG
Department of Business Administration
Hannam University
133 Ojung-dong, Daeduk-gu, Daejeon 306-791
SOUTH KOREA
wonsyang@hnu.kr

TAE-SUNG KIM
Department of Management Information Systems
Chungbuk National University
12 Gaeshin-dong, Heungduk-gu, Cheongju, Chungbuk 361-763
SOUTH KOREA
kimts@chungbuk.ac.kr

Abstract: Since Gordon and Loeb (2002) considered the vulnerability of information to determine the optimal amount of security investment, many researchers has studied security investment decision-making. This article categorizes the types of damage derived from threats, and introduces a stochastic model to incorporate the properties of each type of damage. The results of the stochastic analysis can be used to determine the optimal investment portfolio.

Key-Words: information security investment, optimal investment portfolio, stochastic model

1 Introduction

Since Gordon and Loeb (2002) considered the vulnerability of information to determine the optimal amount of security investment, many researchers have studied on the security investment decision-making (e.g., Tatsumi and Goto (2009) and Huang and Goo (2009)). Threats to information assets incur various types of damages including data loss, hardware replacement and/or repair, and deterioration of system performance. In order to make reasonable decisions regarding information security investment, it is required to consider the economic impact of each type of damage to the information system management.

We consider an information system with three types of threats. First, type-1 threats remove data that the system currently processes. This data become lost. Second, type-2 threats damage hardware and a portion of data stored in hardware. The system requires repairing or replacing the hardware and recovering the damaged data. Finally, type-3 threats deteriorate the system performance, that is, the service rate or the processing speed. The system needs to manage the system performance in order to provide quality service to customers. We first present the stochastic models that describe the above system. Second, we present a financial analysis, which results in the Net Present Value (NPV), presenting the costs and benefits of the system. First, we use the notion of negative customer to describe type-1 threat. Negative customers remove works in the system, which matches the type-1 threat. Queues with negative customers have been studied extensively (See Park et al. (2010) and references therein to review negative customers.) Next, we apply the research of system deterioration to model type-3 threat. We use the model in Yang et al. (2009) and describe the system deterioration by threats and preventive maintenance. The system is monitored continuously and repaired whenever its performance is lower than a predetermined level. Finally, we model type-2 threat using stochastic process.

The system has the following financial structure. The system earns revenue by processing data. The costs consist of the loss cost for the removed data currently being processed, the repair cost for damaged hardware, the recovery cost of the damaged data in hardware, the repair cost for the system maintenance, the holding cost of the system operation, and the security investment.

The rest of this article is structured as follows. In Section 2, we describe the model with some notations. Stochastic analysis on the model is presented in Section 3. The Net Present Value

(NPV) that considers the revenue of the information system, various costs derived from threats, and security investment in order to prevent damage from threats is also presented in Section 3.

2 Model Description

There are M security portfolios. Let PF_m denote the m th portfolio. PF_0 represents the current security level. The security level becomes higher as m increases.

The data that an information system handles, for example, banking and shopping, arrive according to a Poisson process with rate λ. The server has finite states $0,1,\cdots,\beta$, which represent the processing conditions of the system. The processing times are independent exponential random variables with rate μ_k, where k represents the system state. The states are ordered according to the relative degree of deterioration of the system. That is, $\mu_i < \mu_j$ for $i > j$. The system processes the data and stores them in hardware. It is assumed that there is no data at the initiation of the system operation.

Threats are classified into three types according to the damage: First, data that the system processes, including waiting data, are lost by type-1 threats. Let d_k denote the loss probability of the number of data that are lost. It is assumed that the loss probability follows a geometric distribution, such as $d_k = d(1-d)^{k-1}$, $k = 1,2,\cdots$. Second, type-2 threats break down hardware and damage the data that are stored in hardware. The ratio of damaged data is f among the total data. It is assumed that hardware is repaired and data are recovered instantaneously. Third, type-3 threats deteriorate the system, that is, increases the system state by k with probability g_k. Type-i threats occur according to a Poisson process with rate ω_m^i in PF_m. Note that $\omega_k^i < \omega_l^i$ for $k > l$.

We consider a preventive maintenance policy in order to operate the system stable. The system is repaired at or above state α, which we call maintenance level. The repair time is exponentially distributed with rate $1/\delta$. It is assumed that threats do not occur in the system during a repair.

The system earns revenue p per data. The system costs consist of as follows: The loss cost c_L per data, the repair cost of damaged hardware c_W per repair, the recovery cost of damaged data c_D per data, and the holding cost c_H per unit time and data. The system repair cost is c_R^α per repair with a maintenance level of α. The more the system deteriorates, the more resources are needed for repair. Therefore, it is assumed that $c_R^i \geq c_R^j$ for the maintenance level $i > j$. The security investment cost is c_P^m in PF_m. It is assumed that $c_P^i > c_P^j$ for the portfolio $i > j$. The investment occurs at time 0. Finally, we denote the unit fiscal period and the interest rate as τ and θ, respectively.

3 Stochastic and Financial Analysis

First, we analyze the number of data that the system processes by using the Markov Chain. Let (i, j) denote the state of a Markov chain. The notation i represents the number of data in the system and j stands for the system state, for $i = 0,1,\cdots$ and $j = 0,1,\cdots,\beta$. Arranging the states in a lexicographic order gives the following matrix structure:.

$$Q = \begin{pmatrix} B_0 & A_0 & & \\ B_1 & A_1 & A_0 & \\ B_2 & A_2 & A_1 & \ddots \\ B_3 & A_3 & A_2 & \ddots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \quad (1)$$

The matrices A_k and B_k in (1) are the square matrices of the size $(\beta+1)$. The elements of the matrices are shown in Appendix.

Let x_{ij} be the steady-state probability that the Markov chain is in state (i, j). Let π_j be the steady-state probability that the system is in state j. Let us define

$$\boldsymbol{\pi} = (\pi_0,\cdots,\pi_\beta),\ \mathbf{x}_i = (x_{i0},\cdots,x_{i\beta}),\ i = 0,1,\cdots,$$
$$\boldsymbol{\mu} = (\mu_0,\cdots,\mu_\beta).$$

Applying the matrix geometric method (Neuts, 1981) to (1) results in

$$\mathbf{x}_k = \boldsymbol{\pi}(I-R)R^k,\ k = 1,2,\cdots. \quad (2)$$

ISBN: 978-1-61804-150-0

where $\boldsymbol{\pi}$ is the steady-state probability of the Markov chain with the transition rate matrix A.
Let N_m denote the average number of data that the system processes in PF_m. Using (2) gives

$$N_m = \sum_{k=1}^{\infty} k\mathbf{x}_k\mathbf{e} = \boldsymbol{\pi}R(I-R)^{-1}\mathbf{e}. \quad (3)$$

Let us define the throughput and the loss rate as the number of data processed successfully and lost by the type-1 threat, respectively, per unit time. Let Ψ_m and Ω_m denote the throughput and the loss rate in PF_m. Then, we have

$$\Psi_m = \boldsymbol{\pi}\boldsymbol{\mu} = \sum_{j=0}^{\beta} \pi_j \mu_j\,, \quad (4)$$

$$\Omega_m = \frac{\omega_m^1}{d}\boldsymbol{\pi}V\mathbf{e} = \left(\omega_m^1 \sum_{j=0}^{\alpha-1} \pi_j\right)\left(\frac{1}{d}\right). \quad (5)$$

Suppose that the system does not transit to different states when the system is at, or above, state α. In this case, the system behaves stochastically governed by the absorbing Markov chain with the following transition rate matrix $\tilde{A}$.

$$\tilde{A} = \begin{bmatrix} G & \overline{G} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}. \quad (6)$$

The elements of G is given by

$$G = \begin{bmatrix} -\omega_m^3 & \omega_m^3 g_1 & \omega_m^3 g_2 & \cdots & \omega_m^3 g_{\alpha-1} \\ & -\omega_m^3 & \omega_m^3 g_1 & \cdots & \omega_m^3 g_{\alpha-2} \\ & & \ddots & \ddots & \vdots \\ & & & -\omega_m^3 & \omega_m^3 g_1 \\ & & & & -\omega_m^3 \end{bmatrix}.$$

Let Γ_m^1 be the time interval from the point when the system state is 0 to the point when the system repair begins in PF_m. Let Γ_m^2 be the time interval form the initiation to the end of the system repair. Let $\Gamma_m = \Gamma_m^1 + \Gamma_m^2$. Γ_m^1 is equivalent to the absorption time of the Markov chain with the transition rate matrix of (6). Then, we have $E[\Gamma_m^1] = -\mathbf{q}_0 G^{-1}\mathbf{e}$. The average repair time is $E[\Gamma_m^2] = 1/\delta$. This gives

$$E[\Gamma_m] = -\mathbf{q}_0 G^{-1}\mathbf{e} + 1/\delta\,. \quad (7)$$

where $\mathbf{q}_0$ is the column vector of size α and $\mathbf{q}_0 = (1,0,\cdots,0)$.

Let $\Lambda(t)$ denote the average number of data that have been stored in hardware by time t. The data is stored at the rate of Ψ_m in (4). Then, the average number of data in hardware by the j th fiscal period is

$$\Lambda((j-1)\tau) = \Psi_m(j-1)\tau\,. \quad (8)$$

Let Z_k be the point that the k th type-2 threat occurs during $(0,\tau)$. Let Y_k be the interval from the $(k-1)$ th to the k th occurrence point of the type-2 threat. Note that $Z_k = Y_1 + \cdots + Y_k$ and $E[Y_k] = 1/\omega_m^2$. Then, the average number of data stored by Z_k is given by

$$\Lambda(Z_k) = \Psi_m E(Y_1 + \cdots + Y_k) = \frac{\Psi_m}{\omega_m^2}k\,. \quad (9)$$

Let $H_k^{(j)}$ be the average number of data that are damaged by the k th type-2 threat and also recovered during the j th fiscal period $[(j-1)\tau, j\tau]$. The portion of f is damaged among the data in hardware. Then, we have

$$H_k^{(j)} = f\Psi_m\left\{(j-1)\tau + \frac{k}{\omega_m^2}\right\}. \quad (10)$$

Let a_k be the probability that k type-2 threats occur during $[(j-1)\tau, j\tau]$. Type-2 threats occur according to a Poisson process. Thus, we have

$$a_k = \frac{e^{-\omega_m^2\tau}(\omega_m^2\tau)^k}{k!}\,.$$

Let F_m^j be the average number of data recovered during $[(j-1)\tau, j\tau]$ in PF_m. Using (10) gives

ISBN: 978-1-61804-150-0

$$F_m^j = \sum_{k=1}^{\infty} \left\{ H_1^{(j)} + \cdots H_k^{(j)} \right\} a_k$$
$$= \int \Psi_m \left\{ (j-1)\tau(1-a_0) + \frac{1}{\omega_m^2} \sum_{k=1}^{\infty} (1+\cdots+k) a_k \right\}$$

where $a_0 = e^{-\omega_m^2 \tau}$.

Let $P(m, \alpha, y)$ be the NPV at the end of the fiscal period y with a maintenance level α in PF_m. Considering the revenue and cost structure gives the following NPV:

$$P(m,\alpha,y) = -c_P^m + \sum_{j=1}^{y} \frac{L(m,\alpha)\tau - F_m^j c_D}{(1+\theta)^j}$$

$$L(m,\alpha) = p\Psi_m - c_L \Omega_m - c_W \omega_m^2 - \frac{c_R^\alpha}{E[\Gamma_m]} - c_H N_m$$

4 Conclusion

We evaluate information security portfolios considering types of damages: threats which remove data; threats which damage hardware and a portion of data in hardware; and threats which deteriorate systems performance. From the limited availability of data in this paper, empirical or numerical verification has not performed. Only if we obtain data, we can estimate all the parameters (and distributions of parameters) to evaluate information security investment portfolios in order to protect information systems from possible security threats. The model presented in this article can be widely used for evaluating information security investment decisions not only for e-government services but also for private organizations.

References:

[1] Gordon, L.A. and M.P. Loeb, "The economics of information security investment," ACM Transactions on Information and System Security, Vol. 5, pp. 438-457, 2002.

[2] Huang, C.D. and J.H. Goo, "Investment decision on information system security: A scenario approach," Americas Conference on Information Systems 2009, August 2009.

[3] Neuts, M. F., Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach, The Johns Hopkins University Press, Baltimore, 1981.

[4] Park, H.M., W.S. Yang and K.C. Chae, "Analysis of the GI/Geo/1 Queue with Disasters," Stochastic Analysis and Applications, Vol. 28, pp. 44-53, 2010.

[5] Tatsumi, K. and M. Goto, "Optimal timing of information security investment: A real options approach," Workshop on the Economics of Information Security 2009, June 2009.

[6] Yang, W.S., D.E. Lim and K.C. Chae, "Maintenance of deteriorating single server queues with random shocks," Computers and Industrial Engineering, Vol. 57, pp. 1404-1406, 2009.

Appendix : the matrices A_k and B_k

For PF_m and $k = 0,1,\cdots$, we have

$$A_0 = \lambda I,\ A_2 = U + \omega_m^1 d_1 V = U + \omega_m^1 dV,$$
$$A_k = \omega_m^1 d_{k-1} V = \omega_m^1 d(1-d)^{k-2} V,\ k = 3,4,\cdots.$$

where I is the identity matrix of size $(\beta+1)$. U and V are the square matrices of size $(\beta+1)$ and their elements are given by

$$U = \begin{pmatrix} \mu_0 & & & & \\ & \mu_1 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \mu_\beta \end{pmatrix},\ V = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & \ddots \end{pmatrix}.$$

The matrix A_1 is given by

$$A_1 = \begin{pmatrix} -(\lambda+\mu_0+\omega_m^1+\omega_m^3) & \omega_m^3 g_1 & \cdots & \omega_m^3 g_\alpha & \cdots & \omega_m^3 \bar{g}_\beta \\ 0 & \ddots & \omega_m^3 g_1 & \cdots & \cdots & \omega_m^3 \bar{g}_{\beta-1} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \delta & \vdots & \ddots & -(\lambda+\mu_\alpha+\delta) & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \delta & 0 & \vdots & 0 & \ddots & -(\lambda+\mu_\beta+\delta) \end{pmatrix}.$$

where $\bar{g}_j = \sum_{l=j}^{\infty} g_l$. For $k = 1,2,\cdots$, let us define $\bar{d}_k$ as follows: $\bar{d}_k = \sum_{j=k}^{\infty} d_j = (1-d)^{k-1}$. Then the matrices B_k, for $k = 1,2,\cdots$, are as follows:

$$B_0 = A_1 + B_1 = A_1 + A_2 + \cdots,$$
$$B_1 = U + \omega_m^1 \bar{d}_1 V = U + \omega_m^1 V,$$
$$B_k = \omega_m^1 \bar{d}_k V = \omega_m^1 (1-d)^{k-1} V,\ k = 2,\cdots.$$