Scientific Research

Stochastic Design of Enhanced Network Management Architecture and Algorithmic Implementations

Song-Kyoo Kim

W. CySip Graduate School of Business, Asian Institute of Management, Makati, Philippines
Email: SKim@aim.edu

ABSTRACT

The paper is focused on available server management in Internet connected network environments. The local backup servers are hooked up by LAN and replace broken main server immediately and several different types of backup servers are also considered. The remote backup servers are hooked up by VPN (Virtual Private Network) with high-speed optical network. A Virtual Private Network (VPN) is a way to use a public network infrastructure and hooks up long-distance servers within a single network infrastructure. The remote backup servers also replace broken main severs immediately under the different conditions with local backups. When the system performs a mandatory routine maintenance of main and local backup servers, auxiliary servers from other location are being used for backups during idle periods. Analytically tractable results are obtained by using several mathematical techniques and the results are demonstrated in the framework of optimized networked server allocation problems. The operational workflow give the guidelines for the actual implementations.

1. Introduction

In light of the recent acts of terrorism and cyberterrorism, it becomes imperative not only to provide a network security (that has never been a full-proof), but to offer a paradigm of a "network security system" which can be applied to networking for the business continuity such as stock market, postal offices, nuclear power plants, and government offices. Availability [1] of networked servers is a major issue in security respects because of rapid growth of Internet. The paper is focused on enhancement of network availability to support more reliable services. Two types of backup servers are considered. Local backup servers are located in same area and hooked up via LAN (Local Area Network) and he remote backup servers are geometrically separated with main servers but remote backups are hooked up by a Virtual Private Network (VPN) with high-speed optical Internet. A VPN is networking between remote servers and clients via using a public telecommunication infrastructure with secure access to their organization's network. Unlike an expensive system of owned or leased lines, a VPN can provide the organization with the same capabilities, but at a much lower cost. The m main working servers with the $w+1$ local backups and S remote backup servers are geometrically separated with main servers.

The remote (backup) servers are hooked up by a Virtual Private Network (VPN) and can be used during the maintenance of internal backup servers or absence of the repair facility (see **Figure 1**). The number of the remote backups has the control level N within the total number (S) of remote backups. Unlike previous research from author [2], N-policy is applied to restrict the quantity of external resources. A VPN is networking between remote servers and client via using either a public telecommunication infrastructure such as Internet with secure access to private network and secured network such as military system.

In this article we study a class of closed queueing systems with the initial quantity of m main unreliable machines, $w+1$ reserve machines and S auxiliary reserve machines, also called "super-reserve" machines [3-4]. Main working machines are subject to "exponential failures" and their repairs are rendered (in the FIFO order) by a single repair facility (referred to as the repairman) with generally distributed repair times or replacement times to exchange as new machines. As soon as a main working machine breaks down, it is immediately replaced by a reserve machine if available. The total quantity of working machines must not exceed m. Occasionally, the group of reserve machines is blocked for the sake of some routine maintenance, and during this
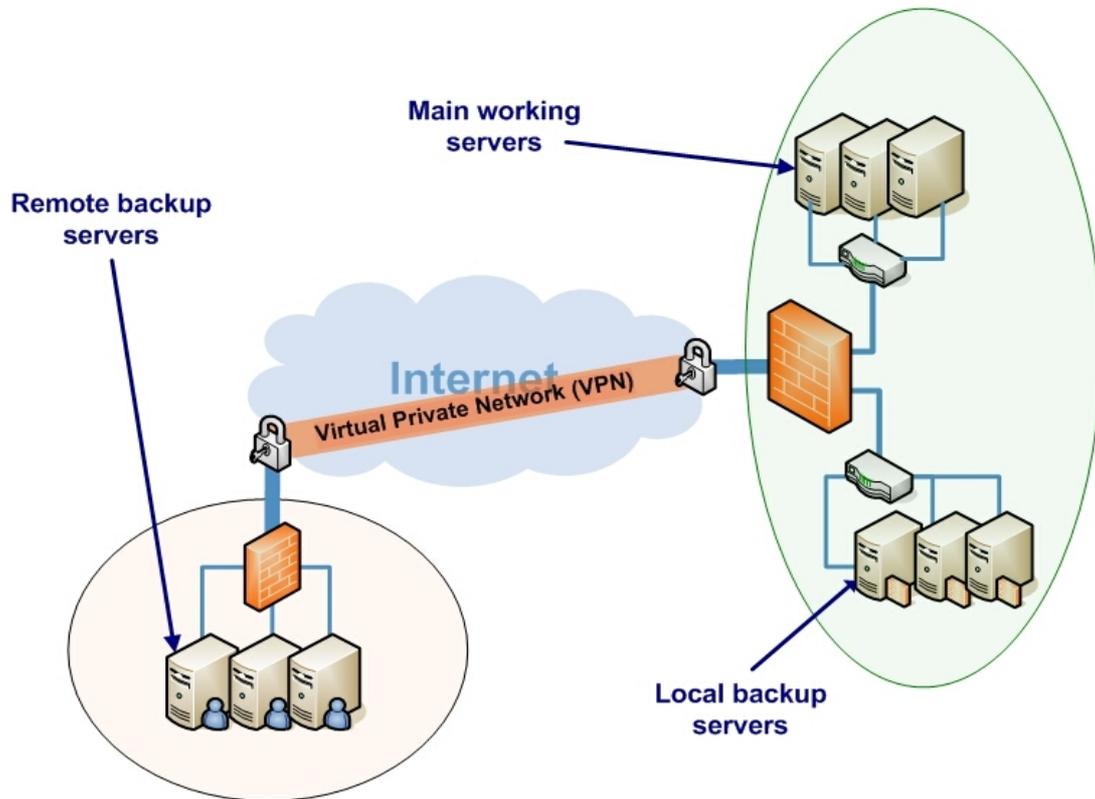
Figure 1. Mixture of the local and the remote backups.

period of time, super-reserve machines take over the duties of reserve machines. The super-reserve facility is "activated" whenever the main and reserve facilities combined are restored to its original quantity $m + w + 1$, and then the system "regenerates". While all main machines keep on working, in the event of failures, the system turns to super-reserve facility, and the repairman is unavailable.

Defective machines are replaced by auxiliary reserve machines whose total number is S. However, the system tries not exhaust this quantity and sets up a smaller control number N. During this period of time, the system is observed only upon some random epochs of time, while dropped machines line up in the "waiting room". If at one of these observation epochs, the number of defective machines reaches or exceeds N (after some delay), the repairman returns to his duties, a busy period begins, and thereby the busy cycle continues. This is a more realistic scenario of a reliability system that functions under restricted observations, at least during its maintenance periods. The scenario can be directly applied for network management. The control integer variable N (less than or equal to S), whose value, among other parameters, is determined in the framework of a comprehensive optimization. Operational workflow gives the implementation guidelines for network management based on the mathematical results. The mathematical val-

ues are the initial conditions for network management operations and the detailed workflow will be explained in this paper.

2. Mathematical Design for Enhance Network Architecture

The Duality Principle [5] is applied and it includes another reliability model, which is more simple than the main model (*i.e.*, Model 1) and to which we will refer as to Model 2. Model 2 is similar to Model 1, except that it does not have the super-reserve facility and idle periods. Besides, the total number of reserve machines is w (*i.e.* less by one than in Model 1. We rather associate it with repairman's vacations, which are distributed as regular repairs. However, upon his return, the repairman brings a brand new machine, which replaces any one that breaks down during his vacation trip if any such available. Otherwise, the new machine he brings in substitutes any other machine and in both cases the old machine is disposed. Model 2 is directly connected with yet another model, which we will call Model 3. Model 3 is a multi-channel queueing system, with buffer of capacity w, and state dependent arrival process, in notation, $(G_1, G_2)/M/m/w$.

Let $\tau_0 (= 0), \tau_1, \cdots$ be the successive moments of repair completions and let $R_1, R_2, \cdots$ be the successive

repair durations all during a busy period. (For brevity of notation, we use $R_1, R_2, \cdots$ as generic random variables for every busy period.) The random variables $\{R_n\}$ are iid with a common probability distribution function [6]

$$A_0(x) := P\{R_n \le x\}, x \ge 0, \qquad (2.1)$$

and mean $a_0 = \mathbb{E}[R_n] < \infty$. Each of the main machines breaks down independently of each other and of repairs, and according to the exponential distribution with parameter $0 < \mu < \infty$. Notice that $\tau_n - \tau_{n-1}$ need not equal R_n, unless the corresponding repair belongs to a busy period. The prebusy period is included in the busy period; the reason for distinguishing this time from the rest of the busy period is for the descriptional convenience and for below arguments regarding the duality principle. We interpret the entire prebusy period as a part of state dependent service with the first service initiating a busy period distributed as the convolution

$$A(x) = \hat{B} \times A_0(x), x \ge 0, \qquad (2.2)$$

where $\hat{B}(x)$ denotes the PDF of the random variable $\tilde{B}_v$. If at time $\tau_n +$ (immediately after the nth repair completion), the total quantity of intact (*i.e.* main working and reserve machines) is less than $m + w + 1$, the busy period goes on.

Model 3 describes the number of customers in a $(G_1, G_2)/M/m/w$ queueing system with state dependent arrival stream. More specifically, it is like a multi-channel queue $GI/M/m/w$ (of Takacs [7]), except for the input is not a "general independent", but it "varies" dependent on the queue length. If upon any arrival, the total number of customers (including those in service) are less than $m + w$, the PDF of the next inter-arrival time is $A_0(x)$. Otherwise, the customers gets lost and the next inter-arrival time is distributed as $A(x)$ of (2.2). While Models 2 and 3 seem to be identical, we call them *stochastically congruent*.

Let $\pi_k^1 := \lim_{t\to\infty} P^x\{Z_t^1 = k\}, k = 0, 1, \cdots, m+w$, be the limiting probabilities of the process Z_t^1. These probabilities exist under the same conditions as those for the embedded process [2].

$$\pi_{m+w+1}^1 = \lim_{t\to\infty} P^x\{Z_t^1 = m+w+1\} = \frac{\xi P_{m+w}}{a_0 + (\xi + \varepsilon) P_{m+w}} \qquad (2.3)$$

and

$$\pi_k^1 = (1 - \pi_{m+w+1}^1)\pi_k, k = 0, 1, \cdots, m+w \qquad (2.4)$$

where

$$\xi = \frac{1}{m\mu} + \mathbb{E}[\tau_v],$$

and

$$\epsilon = \mathbb{E}[B\!\sim_v].$$

And $\pi_k^1 := \lim_{t\to\infty} P^x\{Z_t^1 = k\}$ is subject to our further consideration.

In the other hand, this model is controlled by the so-called "first excess level process" from fluctuation theory. This is a marked three-variate point renewal process with all dependent components. This process by itself can be applied the practical applications such as router design 8. The process will be "terminated" at some of the random observation times when one of its "active" components crosses N, and because its value can be of any magnitude with positive probability, the first excess level will be curtailed to its maximal number S should it formally exceed S. The vacation period ends and the repairman resumes his usual duty. The period of time from t_{n-1} until t_n may or may not include a vacation period and we therefore call it the nth service cycle. During repairman's vacation period, all $w+1$ reserve machines are blocked and the main working facility is backed up by super-reserve machines, which the system "borrows" from a source limited to S units. While all S of them are available, the system attempts to utilize not all super-reserve machines. Namely, it sets up a threshold $N (\le S)$, a specific reference number (to be optimized), the system tries not to exceed.

It is assumed that from the beginning of a vacation period, the status of the system is observed upon some random epochs of time. To simplify notation and without loss of generality we will formalize this process on the first service cycle. Suppose that at t_0, when all of $m + w + 1$ machines become intact, the repairman leaves the main reserve facility, and the system is observed upon the times $\tau_1, \tau_2, \cdots$. We will begin with $\mathbb{E}[\tau_v]$ which is the average period of using the super-reserve machines. Let us assume that the random variables $\tau_{k+1} - \tau_k, k = 0, 1, \cdots$ are exponentially distributed, with common mean $\tilde{\delta}$. By the theorem by author [3-4], we have

$$\mathbb{E}[\tau_v] = \tilde{\delta} + \frac{N}{m\mu}.$$

Now, we turn to $\mathbb{E}[\tilde{\beta}_v]$ that is the average number of super-reserve machine usage:

$$\mathbb{E}[\tilde{\beta}_v] = \tilde{\rho}\tilde{\delta}\lambda\left(1 + \frac{N}{\tilde{\delta}\lambda} - \Delta^{S-N}(S+1) + \Delta^{S-N}S\left(1 + \frac{1}{\tilde{\delta}\lambda}\right)\right)$$

where $\tilde{\rho}$ is the average repair time for single machine. Since $1 + 1/\tilde{\delta}\lambda = 1/\Delta$, we get

$$\mathbb{E}[\tilde{B}_v] = \tilde{\rho}\tilde{\delta}\lambda\left(1 + \frac{N}{\tilde{\delta}\lambda} - \Delta^{S-N}\right).$$

where

$$\tilde{\delta} = \frac{\partial}{\partial\theta}\mathbb{E}[e^{-\vartheta\tau_1}]\Big|_{\theta=0}, \lambda := m\mu.$$

Model 3, as mentioned, is the $(G_1, G_2)/M/m/w$ (multi-channel) queue with state dependent arrivals, m parallel channels, and a buffer or waiting room of capacity w [7]. A customer enters a free channel available with his service demand distributed exponentially with parameter μ. Model 2, as we see it, is congruent to Model 3, while Model 1 is dual with Model 2 (See Dshalalow [5] and Kim [4]). The stationary probabilities $\mathbf{P} = (P_0, P_1, \cdots, P_{m+w})$ for the embedded process are known to satisfy the following formulas:

$$A^{-1} P_k = \begin{cases} \sum_{r=k}^{m-1} (-1)^{r-k} \binom{r}{k} U_k, & k = 0, \cdots, m-1, \\ Q_{m+w-k}, & k = m, \cdots, m+w, \end{cases}$$

where

$$A^{-1} = U_0 + \sum_{j=0}^{w} Q_j, Q(z)$$
$$= Q_0 \cdot \frac{\alpha^0 (m\mu(1-z)) - \alpha (m\mu(1-z)) z}{\alpha^0 (m\mu(1-z)) - z}, \quad (2.7)$$

$$\alpha^0 (\theta) = \int_0^\infty e^{-\theta u} A_0(\mathrm{d}u), \alpha(\theta) = \int_0^\infty e^{-\theta u} A(\mathrm{d}u),$$

$$U_r = a_n^0 \sum_{r=n+1}^{m} \frac{W_r}{a_r \left(1 - \dfrac{1}{a\mu r}\right)}, \alpha_r = \alpha(r\mu), \alpha_r^0 = \alpha^0(r\mu),$$

$$a_s := \begin{cases} \prod_{i=1}^{s} \dfrac{\alpha_i}{1 - \alpha_i}, & s = 1, 2, \cdots, \\ 1, & s = 0, \end{cases}$$

$$a_s^0 := \begin{cases} \prod_{i=1}^{s} \dfrac{\alpha_i^0}{1 - \alpha_i^0}, & s = 1, 2, \cdots, \\ 1, & s = 0, \end{cases}$$

$$W_r = \binom{m}{r} Q_{w+1} \alpha_m^0 - \left[\sum_{j=0}^{w-1} Q_{w-j} S_{jr}^0 + S_{w-1,r} \right],$$
$$r = 0, 1, \cdots, m,$$

$$S_{jr}^0 = \begin{cases} \binom{m}{r} \left[\alpha_r^0 \left(\dfrac{m}{m-r}\right)^{j+1} - \sum_{i=0}^{j+1} \left(\dfrac{m}{m-r}\right)^{j+1-i} \beta_i^0 \right], \\ \qquad\qquad\qquad r = 0, 1, \cdots, m-1, \\ 0, \qquad\qquad\qquad r = m, \end{cases}$$

$$S_{w-1,r} = \begin{cases} \binom{m}{r} \left[\alpha_r \left(\dfrac{m}{m-r}\right)^{w} - \sum_{i=0}^{w} \left(\dfrac{m}{m-r}\right)^{w-i} \beta_i \right], \\ \qquad\qquad\qquad r = 0, 1, \cdots, m-1 \\ 0, \qquad\qquad\qquad r = m, \end{cases}$$

$$\beta_k^0 = \int_{\mathbb{R}_+} e^{-m\mu x} \cdot \frac{(m\mu)^k}{k!} \mathrm{d}A(x), k = 0, \cdots, m-1,$$

$$\beta_k = \int_{\mathbb{R}_+} e^{-m\mu x} \cdot \frac{(m\mu)^k}{k!} \mathrm{d}A(x), k = m,$$

and $Q(z)$ of (4.1) is the generating function, convergent in the open disc centered at zero. By using the Kolmogorov differential equation and the semi-regenerative techniques [2-4,8-9], this system has been solved by Dshalalow [6]. The limiting distribution $\pi = \{\pi_0, \pi_1, \cdots, \pi_{m+w}\}$ is:

$$\begin{cases} \pi_0 = 1 - \sum_{n=1}^{m+w} \pi_n, k = 0, \\ \pi_k = \dfrac{P_{k-1}}{k\mu \mathbf{PA}}, k = 1, \cdots, m-1, \\ \pi_k = \dfrac{P_{k-1}}{m\mu \mathbf{PA}}, k = m, \cdots, m+w \end{cases} \quad (2.8)$$

where

$$\mathbf{PA} = a_0 (P_0 + \cdots + P_{m+w-1}) + a P_{m+w}$$

For the process Z_t^1, the corresponding formulas yield

$$\pi_k^1 = \frac{a_0 + \epsilon P_{m+w}}{a_0 + (\xi + \epsilon) P_{m+w}} \pi_k, k = 0, \cdots, m+w, \quad (2.9)$$

$$\pi_{m+w+1}^1 = \frac{\xi P_{m+w}}{a_0 + (\xi + \epsilon) P_{m+w}}, \quad (2.10)$$

along with (2.8).

3. Networked Server with Controlled Backup Optimization

The stochastic optimization techniques are used for the sample illustration of the optimization and the stochastic optimization techniques by itself can be applied to real-world problems such as computer-networking, human resources and manufacturing process. Let a strategy, say Σ, specify, ahead of the time, a set of acts we impose on the system and the system can be subject to a set C of cost functions. The general formula of stochastic optimization is [2-4]:

$$\phi(\Sigma, C) := \lim_{t \to \infty} \frac{1}{t} \phi(\Sigma, C, t) \quad (3.1)$$

Now we turn to convergence theorems for semi-regenerative, semi-Markov, and Markov renewal processes [10],

$$(\mathrm{i}) \lim_{t \to \infty} \frac{1}{t} R^i(t) = \frac{1}{\mathbf{PM}}$$

$$(\mathrm{ii}) \lim_{t \to \infty} \frac{1}{t} \int_0^t P^i \{Z_s^1 = k\} \mathrm{d}s = \pi_k^1$$

$$(\mathrm{iii}) \lim_{t \to \infty} \frac{1}{t} \int_0^t P^i \{Y_u^1 = k\} \mathrm{d}u = \frac{P_k M_k}{\mathbf{PM}}$$

to arrive at the objective function $\phi(\Sigma, C)$, which gives the total expected rate of all processes over an infinite horizon. As a reasonable performance measure, let us

consider the reliability factor ρ, which represents the probability of the number of intact machines at any moment of time in equilibrium:

$$\rho = \sum_{k=m}^{m+w+1} \pi_k^1. \tag{3.2}$$

This is not only a reliability measure of the system, but it can also serve as a constraint to an optimally functioning the system. We arrive at the following expression for the sample objective function [2]:

$$\begin{aligned} \phi\big(\Sigma(N),C\big) \\ = c \cdot \pi_{m+w+1}^1 + \big(G-B\big)\overline{Z}_\infty^1 \\ + \big(H-G\big)\sum_{k=m+1}^{m+w+1}\big(k-m\big)\pi_k^1 \\ + B\big(m+w+1\big)+r \cdot \frac{1}{PM} \end{aligned} \tag{3.3}$$

Take the total number of main networked servers as 2 and the total number of local backups is 4. We are setting up the maximal availability of remote backup servers to 15. Hence, $m=2, w+1=4$ and $S=15$. Now, we calculate $\phi\big(\Sigma(N),C\big)$ and N_0 that gives a minimum for $\phi\big(\Sigma(N),C\big)$. In other words, the control level N_0 stands for the excess level of remote backup which minimizes the total cost of this system. Below is a plot of $\phi\big(\Sigma(N),C\big)$ for $N \in \{1, 2, \cdots, S(=15)\}$.

Our calculation yields that $N_0 = 9$ for which the minimal cost equals 15.6445. It means that we allocate our internal resources to 2 main (m) and 4 internal $(w+1)$ networked servers and obtain the threshold value $N_0 = 9$ which gives us the decision point that is the number of remote backups which we need from external resources to minimize the cost of the backup system. Using the above example of our model, we arrive at the reliability factor is $\rho = 0.2607$. It tells us that the likelihood of having at least m intact main networked server is 0.2607.

4. Algorithmic Implementations for the Enhanced Network Architecture

The network architecture that has mentioned in the previous is the mathematical and theoretical approach to analyze the stochastic model. The operational method is the guideline for actual implementation. The workflow of operating the enhance network management can be easily adapted for software programming and simulation. All of the mathematical results from the previous sections are applied into the operational method as the initial conditions. The variables need to be defined for using the results from the mathematical model. Number of iteration, number of main and backup servers, the status of the repair facility are some of key factors for implementation. The variables for operations of enhanced network management are as follow:

n : number of iterations
m_n : number of main servers at iteration n
w_n : number of local backup servers at iteration n
S_n : number of remote backup servers at iteration n

$$Q_n := m_n + w_n$$

L_0 : control level of remote backup servers
c : counting number of remote backup usage
G : number of servers that have been fixed within Iterations

Repaireman :

$$= \begin{cases} ON, \text{repair facility is available}\big(\text{working}\big), \\ OFF, \text{repair facility is on vacation.} \end{cases}$$

The values in the mathematical model are applied as the initial conditions in the operational workflow but the notations are different. The delta of notations between mathematical model and operational method is shown in **Table 1**.

The operational workflow can be presented after defining the initial condition (see **Figure 2**) The workflow is the depiction of a sequence of operations for enhanced network management that is focused on service availability.

If the operations is applied in the example case in Section 3, the actual values of the initial condition are given:

$$m_0 = 2, w_0 = 4, S_0 = 15, L_0 = 9$$

based on the delta list (see **Table 1**) The network management based on above operation workflow gives the optimal performance in server availability perspective.

5. Conclusion

In this article theoretical approaches of the network defense model is presented. Unlike simulated model, we can find the explicit formulas that is the key elements of the complex model. In addition, this model can be also applied various real-world applications such as network system design 8 and software architecture [11]. Analytically tractable results are obtained by using a duality principle

Table 1. Delta list of the initial condition.

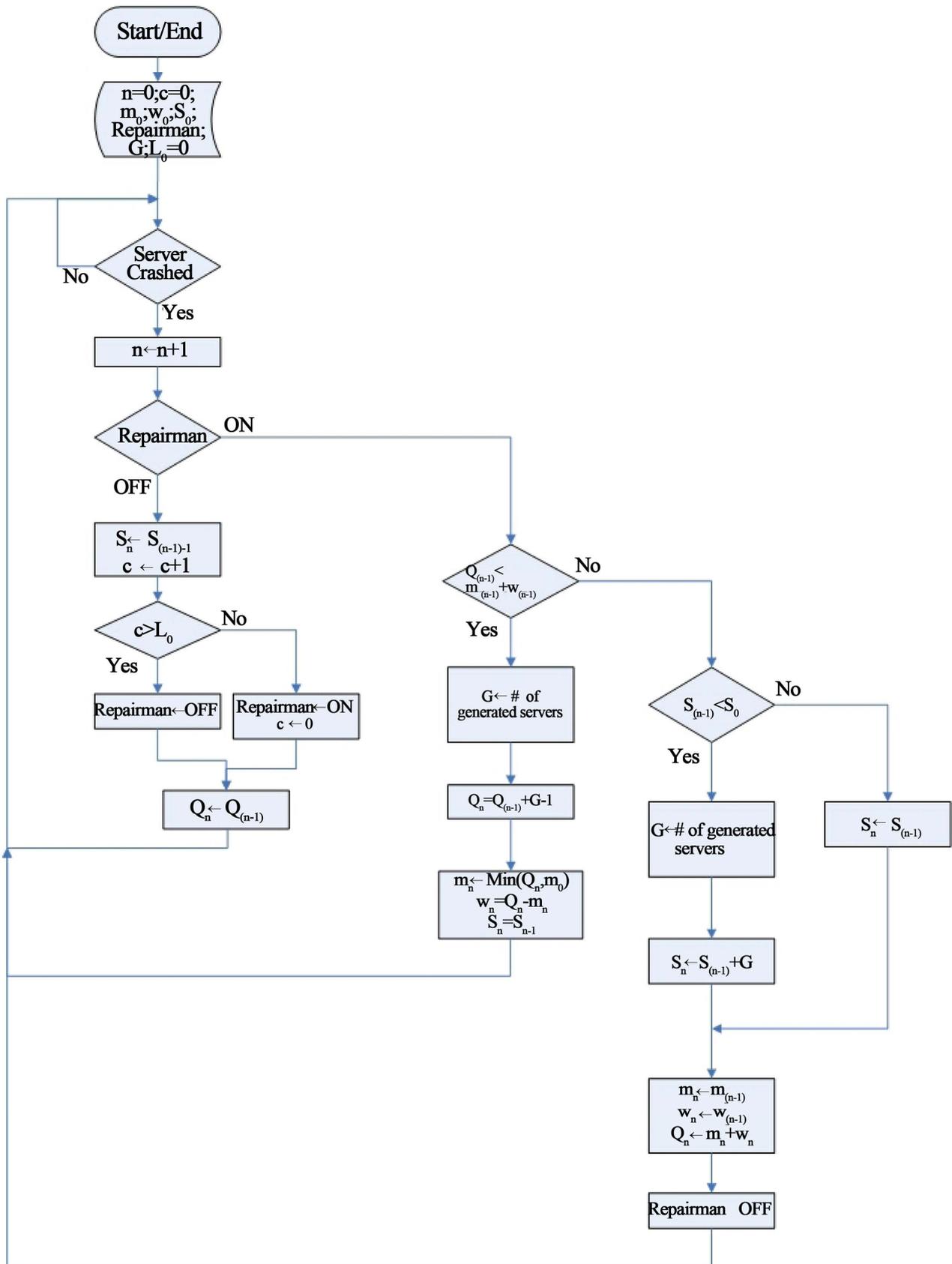| Opt.	Math.	Descriptions
m_0	m	number of main servers
w_0	$w+1$	number of local backup servers
S_0	S	number of remote backup servers
L_0	N_0	first exceed level of remote servers

Figure 2. Operations workflow of enhance network.

(which enables us to treat a more rudimentary system), semi-regenerative analysis, and the theory of fluctuations of multivariate marked renewal processes. The results are applied in the framework of optimization problems.

REFERENCES

[1] D. Russell and G. T. Ganemi Sr., "Computer Security Basics," O' Reilly and Asso. Inc., Sebastopol, 2006.

[2] S.-K. Kim, "Enhanced Networked Server mgt. with Random Remote Backups," *Proceedings of SPIE* 5244, Orlando, 7 September 2003, pp. 106-114. doi:10.1117/12.511412

[3] S.-K. Kim, "Enhanced Management Method of Storage Area Network (SAN) Server with Random Remote Backups," *Mathematical and Computer Modelling*, Vol. 42, No. 9-10, 2005, pp. 947-958. doi:10.1016/j.mcm.2005.06.006

[4] S.-K. Kim, "Enhanced Stochastic Methodology for Combined Architecture of e-Commerce and Security Networks," *Mathematical Problems in Engineering*, Vol. 2009, 2009, Article ID: 691680.

[5] J. H. Dshalalow, "On a Duality Principle in Processes of Servicing Machines with Double Control," *Journal of Applied Mathematics*, Vo. 1, No. 3, 1988, pp. 245-251.

[6] J. H. Dshalalow, "Queueing Systems with State Dependent Parameters," In: J. H. Dshalalow, Ed., *Frontiers in Queueing*, CRC Press, Boca Raton, 1997, pp. 61-116.

[7] L. Takacs, "Some Probability Questions in the Theory of Telephone Traffic," *Magyar Tudományos Akadémia. Matematikai és Fizikai Osztály Közleményei*, Vol. 8, 1958, pp. 155-175.

[8] S.-K. Kim, "Design of Stochastic Hitless-Prediction Router by Using the First Exceed Level Theory," *Mathematical Methods in the Applied Sciences*, Vol. 28, No. 12, 2005, pp. 1481-1490. doi:10.1002/mma.626

[9] J. H. Dshalalow, "On the Multiserver Queue with Finite Waiting Room and Controlled Input," *Advanced Applied Probability*, Vol. 17, No. 2, 1985, pp. 408-423. doi:10.2307/1427148

[10] E. Cinlar, "Introduction to Stochastic Processes," Prentice Hall, Englewood Cliffs, 1975.

[11] S.-K. Kim, "Design of Enhanced Software Protection Architecutre by Using Theory of Inventive Problem Solving," *IEEE Proceedings of IEEM*, Hong Kong, 8-11 December 2009, pp. 978-982.