Journal of the Korean Institute of Illuminating and Electrical Installation Engineers Vol. 23, No.11, pp. 34~43 November 2009 (23-11-5)

Key Distribution Process for Encryption of SCADA Communication using Game Theory applied Multiagent System

Hak-Man Kim* · Dong-Joo Kang**

Abstract

SCADA (Supervisory Control and Data Acquisition) system has been used for remote measurement and control on the critical infrastructures as well as modern industrial facilities. As cyber attacks increase on communication networks, SCADA network has been also exposed to cyber security problems. Especially, SCADA systems of energy industry such as electric power, gas and oil are vulnerable to targeted cyber attack and terrorism. Recently, many research efforts to solve the problems have made progress on SCADA network security. In this paper, flexible key distribution concept is proposed for improving the security of SCADA network using Multiagent System (MAS).

Key Words : SCADA(Supervisory Control and Data Acquisition) System, Cyber Security, Multiagent System(MAS), Game Theory

1. Introduction

SCADA(Supervisory Control and Data Acquisition) system is a system operation with coded signals over communication channels so as to provide control of RTU (Remote Terminal Unit) equipment [1]. Recently Intelligent Electronic Device (IED) which is control unit having communication function with master station is replacing the role of RTU.

SCADA system has been used for remote measurement and control on the critical infrastructures such as electric power, gas and oil as well as modern industrial facilities such as chemical factories, manufacturing facilities. SCADA network has been exposed to cyber security problems with IT advancement and network growth. Especially, SCADA systems of energy industry such as electric power, gas and oil are vulnerable to targeted cyber attack and terrorism. Recently, research efforts to solve the problems have been progressed throughout the world.

System (MAS) is one of artificial intelligence (AI). The common characteristics of MAS are

* Main author : Department of Electrical Engineering, Incheon City College
** Corresponding author : Korea Electro-technology Research Institute
Tel : +82-31-420-6181, Fax : +82-31-420-6189
E-mail : djkang@keri.re.kr
Date of submit : 2009. 8. 18
First assessment : 2009. 8. 19
Completion of assessment : 2009. 9. 15

autonomy, social ability and intelligence. MAS has been researched in security areas as well as various engineering areas [2-5].

In this paper, flexible key distribution scheme is proposed for SCADA network security. Flexible key distribution supports more secure ability than fixed key distribution. MAS is applied for flexible key distribution. For applying MAS, multi agents are defined. Also, their functions, relationship and information flow are introduced.

2. SCADA network security

SCADA system has been used to remote measurement and control for the critical infrastructures as well as modern industrial facilities. Fig. 1 shows the general configuration of SCADA network. The communication system links the control center with IEDs. Common methods of communication include radio, leased line, landline, and digital and analog microwave. More recently analog and digital cellular communication has been introduced. For remote service, satellite communication is sometimes employed. SCADA security in communication typically refers to the ability to perform error correction, rather than authentication or encryption [6].

Several trends have led to an environment in which the security of SCADA system has become more critical. The main trends are as follows [7]:

a) The use of common operating systems, such as Microsoft Window and Unix, in SCADA and control system platforms;
b) The increased use of TCP/IP communications;
c) The demand from corporate users for operational data on a near-real-time basis.

Recently, research efforts to solve the problems have been progressed in SCADA network security. There are many research challenges such as access control, firewall, intrusion detection system, protocol vulnerability assessment, cryptography, key management, devices/OS security and security management for SCADA networks [8].

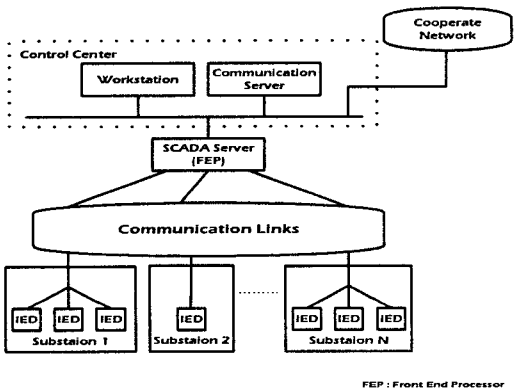

Fig. 1. SCADA network configuration

3. Game theoretic approach to MAS

The concept of "agent" has been also introduced in artificial intelligence when we describe the entity or program which performs some tasks specialized in some field or very complicated to be done instead of a person. Multiagent is a collection of more than one agent and has common characteristics such as autonomy, social ability, and intelligence. Autonomy means agent can judge by itself and perform some tasks based on the judge without the orders from user or other programs. Social ability means the agent's capability to cooperate with other agents to perform some tasks or accomplish some objects. Intelligence means agent performs tasks or accomplish objective according to its own reasoning and judging process not just based on codes or program already made by human.

Relationship between agents is duly defined using game theory. Game theory provides us with

three different kinds of gaming situation classified as cooperative game, non-cooperative game, and negotiation (or bargaining) game. In the same fashion, each game situation has a few sub-models like Cournot model, Bertrand model, Stackelberg game, Nash bargaining game, etc. Fig. 2 shows the relation between payoff superstition and competition intensity in general economics.

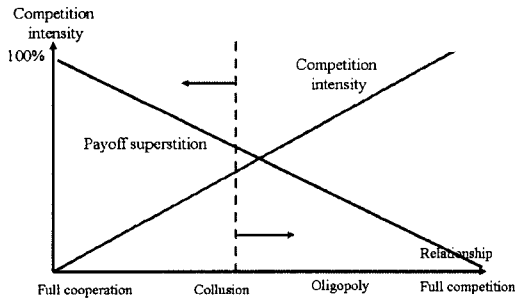

Fig. 2. Relationship modeling based on competition level

A game is specified as n players, their strategies and payoff matrix. If two payoff matrices of two players are the same or proportional, the game of two players is a fully cooperative game. If the payoff matrix is completely different from each other, the game would be an on-cooperative game.

		B's action				B's action		
		R	P	S		R	P	S
	R	0	-2	+1	R	0	+2	-1
A's action	P	+2	0	-1	P	-2	0	+1
	S	-1	+3	0	S	+1	-3	0
		A's payoff				B's payoff		

Fig. 3. Example for payoff matrices of two players

The strategic choice for the relationship between two agents is wholly determined by the judgment of agents. The choices of agents may be different from each other even in same situation dependent on the objective function or the strategic preference of each agent.

4. MAS Applied Key Distribution Process

Encryption strengthens the security by protecting the network from attack and thereby decreasing the vulnerability of network. However the encryption itself is always exposed to the danger of being cracked.

So we change the secret key of encryption periodically and the level of danger would increase as the time duration of key distribution period lasts longer. But if we make the period too short, it could cause the inefficiency and load increase of network and key distribution server. In this aspect we need some kind of key distribution policy to keep or increase the security level of encryption while maintaining the efficiency of the system, and we propose a flexible key distribution scheme based on Multiagent concept and security assessment method.

4.1 Agent based design framework

MAS based system makes it possible to allocate many functions centrally controlled to local agents, which is expected to increase the efficiency and flexibility of the system.

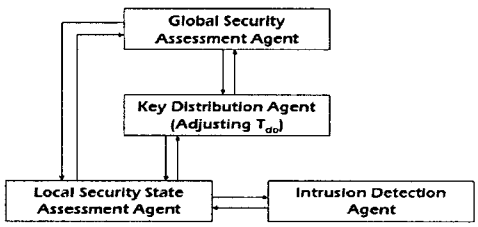

Fig. 4. Multiagent based flexible key distribution

Fig. 4 shows us MAS based key distribution process. T_{dp} stands for key distribution period. IDA (Intrusion etection agent) monitors whether intrusion is or not periodically on a specific node

or an area which it is in charge of. IDA has previously input knowledge on many types of intrusion patterns and has capability of learning and analyzing new patterns of attacking. IDA could reside in each communication node or have control of a local area composed of more than one node. Fig. 5 shows node based IDA. Each IDA stays in its responsible node continuously and watches if there is intrusion at the node.

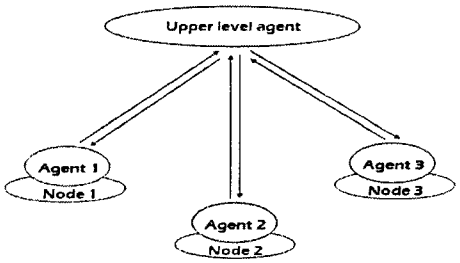

Fig. 5. Node based ID agent

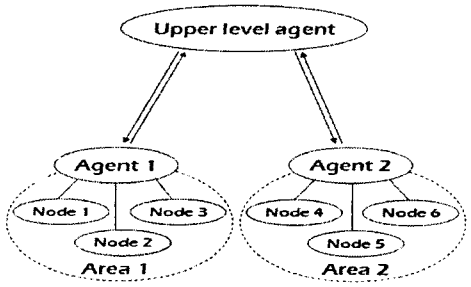

Fig. 6. Area based agent

Fig. 5 shows area based IDA. When there are too many nodes in the system, it is impossible or inefficient to dispatch all agents to all nodes. In this case we could divide the whole system into several areas by grouping those nodes. Agent is in charge of its nodes in its responsible area.

Local Security State Assessment (LSSA) agent performs the analysis on each node or each area, and thereby they check local based security level or vulnerability based on mathematical model using random variable distribution model.

Global Security Assessment (GSA) agent performs the assessment global security state or vulnerability of whole system. When the security key is renewed based on whole system, key distribution (KD) agent communicates with VIA agent to adjust the period of key distribution according to the result of security assessment of GSA agent. When this key distribution is done at each area, key distribution agent communicates with LSSA agent directly to determine the new period of key distribution. It depends on security or key distribution policy, and Fig. 7 shows this communication and calculation flow between agents dependent on the policy.

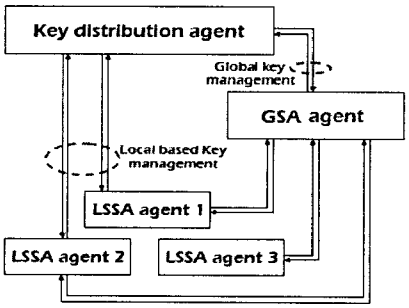

Fig. 7. Communication between agents

4.2 Relationship modeling between agents using game theory

We introduced 4 agents in Fig. 4, which are KDA, IDA, GSA agent, and LSSA agent. Basically all agents in the system are required to cooperate with each other. They have common object to accomplish the good performance and security of the whole system. Therefore the relationship between agents could be modeled using cooperative game theory. However there might be also a conflict between agents like GSA and LSSA agents. GSA covers whole system and is responsible for the reliability of entire system, while LSSA is a local agent in charge of security

and performance monitoring in its node or region. Here we define the reliability as the overall index induced by the security and performance level of the system. When we consider the local and global reliabilities at the same time, it is required to be balanced between local and global aspects. Considering this situation, we could assume the trade-off between local and global reliability as bargaining process between LSSA and GSA agents, and thereby the NBS (Nash bargaining solution) could be an appropriate solution to model the relationship between two agents.

4.2.1 Nash bargaining solution

Nash bargaining solution is used for modeling cooperative games of two players

When there are two bargainers, A and B seeking to split a total value v which they can achieve if and only if they agree on a specific division, they would get their final outcomes x for A and y for B respectively. Because two bargainers are supposed to share the common outcome v, the equation x+y=v should be fulfilled. If no agreement is reached, A will get a and B will get b, each by acting alone or in some other way acting outside of this relationship. Thinking differently, a and b could be considered as the initial cost of investment for obtaining the outputs, x and y, or the opportunity cost of x and y. Here, a and b are named as their backstop payoffs or, in the jargon of the Harvard Negotiation Project, BATNAs (best alternative to a negotiated agreement) [9]. Of course a+b<v should be satisfied. When we assume two bargainers divide the common surplus v with the fraction of α for A and β for B, two bargainer's outcomes x and y could be also expressed respectively as the sum of each BATNA and backstop as follows:

$$x = a + \alpha(\nu - a - b) \rightarrow x - a = \alpha(\nu - a - b) \quad (1)$$

$$y = b + \beta(\nu - a - b) \rightarrow y - b = \beta(\nu - a - b) \quad (2)$$

If we divide (1) by (2), we could get another form of equation as follows:

$$\frac{x-a}{y-b} = \frac{\alpha}{\beta} \quad (3)$$

There are three principles prerequisite to model cooperative game as Nash bargaining solution [10].

a) The outcome should be invariant when the scale of payoffs changes linearly.
b) The outcome should be efficient, which means x+y=v ,therefore there is no unexploited gain.
c) The outcome should be independent from other variables except x and y we are considering.

When three assumptions are fulfilled, the bargaining game can be modeled as following Nash bargaining formula.

$$Max \ (x-a)^{\alpha}(y-b)^{\beta} \quad (4)$$

subject to y = f(x)

Besides three assumptions mentioned above, Nash originally imposed fourth assumption α=β which means both parties share the outcome equally, thereby equation(4) could be re-written on the unique case.

$$Max \ (x-a)^{1/2}(y-b)^{1/2} \quad (5)$$

4.2.2 Application of NBS to relationship model between agents

When we consider Fig. 7, the agents are in cooperative relation and need to interact with each other to increase the reliability of entire system.

The reinforcement of local security generally increases the global security of the entire system. However there is always the problem of limited resources and is required to allocate the limited resources in optimal way.

What is the limited resource in network communication and encryption process? One of the resources would be time. Encryption process causes network traffic increase and thereby communication time delay which result in the reliability degradation. Therefore we need to find a balance between security strength and network performance, which is the role of agents like LSSA or GSA agents in Fig. 7. The second resource would be the communication network shared by many entities or agents. The network is also limited to be fully used by each of all entities on communication.

Considering two aspects mentioned above there are two bargaining dimensions on the problem. One is the bargaining between security and performance. The other is the bargaining between LSSA and GSA agents for optimizing the resource allocation. This concept is illustrated in Fig. 8.

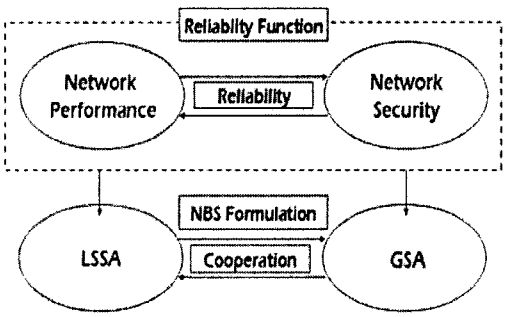

Fig. 8. NBS application to Agents' relationship

Security and performance are co-optimized in the reliability function and thereafter agents adjust for the optimal balance on the entire system aspect.

4.3 Mathematical formulation of reliability function

We already mentioned in previous chapter that there is inversely proportional relationship between security level and system performance. We could strengthen the security level by increasing the frequency of key distribution, while we should endure the degradation of system performance. So we should find out the balance or equilibrium between two factors, but there is a difficulty adjusting two factors because they exist in different dimensions. We need to unify the units of two values. In that aspect two factors are required to be commensurable first of all. We could define a kind of QoS(Quality of Service) function composed of network traffic part and security part, which represents the overall quality of service on communication.

$$QoS = PI + SI \tag{6}$$

Here PI and SI mean performance index and security index respectively. We model (6) as the function of key distribution period, Tdp defined in Fig. 4, and notate it as td from now on.

4.3.1 Performance Index model

PI is calculated based on the time delay caused by encryption process, and the value of PI is distributed between 0 and 1. There could be many causes of communication error or failure, but we confine the cause only to the time delay caused by encryption and decryption process. We could assume a functional relationship between the communication delay and the key distribution period.

If the period gets shorter, communication delay would be longer, because the frequent key distribution increases the network traffic on the

network. We assume the time delay on communication by encryption process δ, and δ can be formulized as the function of td as δ=f(td).

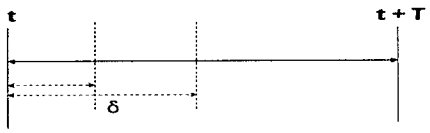

Fig. 9. Time delay by encryption process

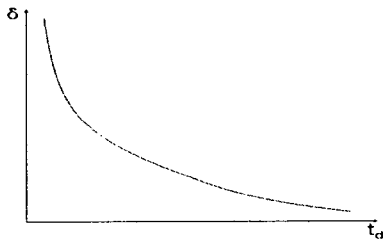

Fig. 10. Functional relationship between δ and td

And δ is inversely proportional to td, and expressed as T/td as shown in Fig. 10. Here, T is the period of SCADA communication. To normalize the value, we correct the formula as (T-δ)/T. When δ reaches T, the performance value converges to 0 which means communication failure. We just assume the functional relationship between δ and td as δ=k/td as shown in Fig. 10. Here k is a constant. Therefore the performance index related to time delay by encryption is formulated as follows.

$$PI_t = \frac{T - k/t_d}{T} \tag{7}$$

4.3.2 Security Index model

In this paper we model the intrusion event and vulnerability index as random variable. We assume the number of cyber intrusion happens with the Poisson process and thereby the time duration between two events happening follows the probability of exponential distribution function. Each cyber intrusion is made by each person therefore it could be defined as a independent event.

Let us assume the Poisson process {N(t), t≥0} which has a Poisson having rate λ(λ>0). We could apply Poisson process to the modeling of real world problem when we fulfill following conditions [10]:

a) N(0)=0;
b) The process has independent increments;
c) The number of events in any interval of length t is Poisson distributed with mean λt. That is, for all s.t. ≥0

$$P\{N(t+s) - N(s) = n\} = e^{-\lambda t}\frac{(\lambda t)^n}{n!},\ n = 0, 1, \ldots.$$

We can consider the number of intrusions or intrusion trials within certain time duration is defined as N(t). λ is average intrusion rate of the Poisson process applied to intrusion number modeling. When N(t) follows the Poisson process, the time interval t follows the following exponential distribution function.

$$P\{T > t\} = e^{-\lambda t} \tag{8}$$

Here P{T>t} means the probability of time interval, T is longer than t from the occurrence of previous intrusion to next one. That means no intrusion occurs [0, t], so we can mark it as P{no intrusion}.

$$SI_t = e^{-\lambda t_d} \tag{9}$$

4.3.3 Reliability function model

Reliability is measured as the QoS defined in equation (6). Based on (7) and (9), reliability function is defined as follows:

$$RI_t = QoS(t_d) = 1 - k/(T \cdot t_d) + e^{-\lambda t_d} \tag{10}$$

Here, RI stands for Reliability Index, and RI is used as a standard for agents to build their security policy, the key distribution period specifically in this paper.

4.4 NBS model for the relationship between LSSA and GSA agents

The result from (10) should be also adjusted by global security policy involving many agents. We defined a GSA agent and multiple LSSA agents in Fig. 11.

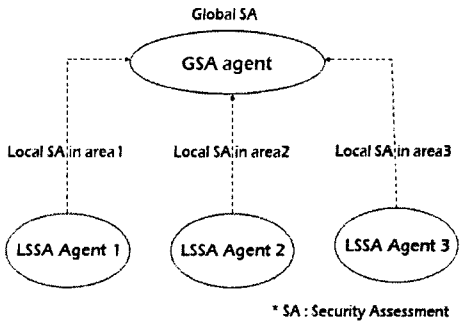

Fig. 11. Global security assessment

There could be many alternative ways to calculate the global security, notated as GS. One of them could be a weighted average sum in linear form as follows:

$$GRI = w_1 LRI_1 + w_2 LRI_2 + \ldots + w_n LRI_n \qquad (11)$$

This function is embedded or programmed in GSA agent. GRI and LRI stands for global reliability index and local reliability index respectively. LRI is calculated by LSSA agent using equation (10). Equation (11) is the one of simple methods for inducing a global reliability index based on linear weighted sum of many local reliability indices. There could be many alternatives for calculating GRI dependent on system characteristics, and even (11) might not be appropriate. We assume (11) as the right method for GRI calculation in this paper and we put off the problem about GRI calculation to future studies.

When we accept the model (11), the following problem is to calculate the weight, wi. Fortunately the SCADA network is the simple radial form having one master station (of sub-SCADA or central SCADA) connected with all RTUs respectively. Considering this aspect we could impose same values on all 'w's and thereby wi=1/n. To apply NBS to this problem, we need to specify two bargaining entities. There are two relationships for NBS application as shown in Fig. 8. One is between PI and SI on the aspect of reliability optimization. The other is between LSSA and GS agents. There is only one strategic variable, key distribution period, td in this paper. Considering the situation LSSA and GSA cannot be the bargaining entities because the td is determined by LSS agent. Therefore, in this paper, we only consider PI and SI to apply NBS to MAS based key distribution process. Two entities outcomes, x and y in the NBS formulation above, would be PI and SI which are quantitative indices of performance and security. Backstops of two bargainers are equally 0.

Using (7) and (9), we can induce the relational equation between PI and SI. (7) and (9) are able to be developed as the form of td = ·.

$$t_d = \frac{k}{T - T \cdot PI_t} \qquad (12)$$

$$t_d = -\frac{\ln SI_t}{\lambda} \qquad (13)$$

By substituting (12) for (13), we induce the relational equation of two variables, PI and SI. SI can be the function of PI.

$$SI_t = e^{-\frac{\lambda k}{T - T \cdot PI}} \qquad (14)$$

Because the necessary conditions of NBS are fulfilled and a=β=1/2, the NBS of this problem is formulized as follows:

$$\underset{PI,SI}{Max}\ PI^{1/2}SI^{1/2} \tag{15}$$

subject to $SI_t = e^{-\frac{\lambda k}{T-T\cdot PI}}$

By solving (13) we could get optimal PI and SI, which also gives us optimal td. So we could decide the distribution period of security key for maximizing SCADA system reliability.

4.5 Application to SCADA network

There are many options to divide and operate the SCADA network for applying security policy like key distribution. For example a region and its nodes within it could compose a unit communication network like sub-SCADA in charge of a LSSA agent as shown Fig. 12. In this case a LSSA agent resides in sub-SCADA server for monitoring and assessment of the security state. And VIA agent in central SCADA server assesses the reliability of whole system by communicating with LSSA agents and calculating GRI based on the LRI data provided from local LSSA agents according to equation (11). Of course there could be also upper level agent staying in SCADA server to manage and control substation level agents.

Once the agents are defined on their role and functions, it is important to model their relationship. We showed one of the methods based on NBS in this paper. MAS based scheme is basically composed of independent agents, which are entities making their own decisions independently. So the key aspect of integrating MAS based system into a single process is to define the relationship between agents. There could be also lots of variables to affect system reliability besides key distribution period. As the number of variables increase the complexity of MAS based model would be complicated. We will consider it in future studies.

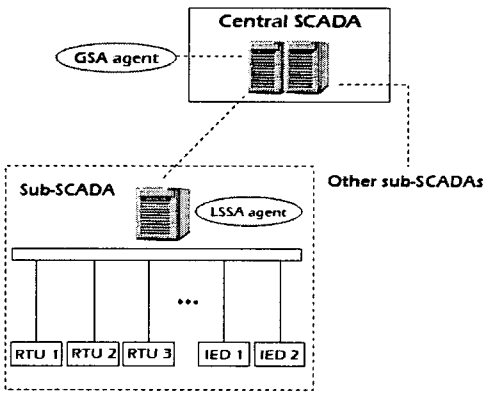

Fig. 12. Application to SCADA network.

5. Conclusion

Cyber security problems of SCADA network of critical infrastructures such as electric power, gas and oil are very important against cyber attack and terrorism. Recently, research efforts to solve the problems have been progressed throughout the world. In this paper, flexible key distribution scheme is suggested for SCADA network using Multiagent System.

In this paper, flexible key management scheme based on MAS is proposed for enhancing security level of SCADA network. And we modeled it using NBS, one of game theories for cooperative game. There are many kinds of alternatives for game theories, which model applied would be dependent network or communication characteristics and objective function. We use the NBS for cooperative game model because we intend to find out a balance between performance and security. So there could be more flexibilities and varieties

for the key management process using game theory applied MAS. Based on these diversities the flexible key distribution based on MAS supports more secure and efficient system than under fixed key distribution. Key contributions of this paper are as follows:

a) Flexible key management for secure SCADA network based on MAS;
b) Definition and function of each agent introduced, and relationship and information flow between agents to form MAS;
c) Security state defined based on mathematical formulation using probability and random variable process;
d) Vulnerability index concept introduced to assess the global level of reliability or security of the system against cyber attack.

More detail and developed research is needed for objective function formulation of each agent and their relationship in both aspects of mathematical model and programming dimension. In addition, many different kinds of approaches could be tried such as ANN (Artificial Neural Network), GA (Genetic Algorithm) to model the agent's intelligence and the relationship between agents. These various methods will be helpful for realizing MAS based system and improve MAS intelligence, which contributes to developing more secure encryption system.

Acknowledgement

This study was supported by Incheon City College in 2009.

References

[1] A.B. Smith, IEEE Std C37.1-1994, IEEE Standard Definition, Specification, and Analysis of Systems Used for Supervisory Control, Data Acquisition, and Automatic Control ,IEEE Power Engineering Society, Sponsored by the Substations Committee, Institute of Electrical and Electronics Engineers, Inc., New York, 1994.
[2] G.G. Helmer, J.S.K. Wong, V. Honavar and L. Miller, "Intelligent Agents for Intrusion Detection", Proc. of the IEEE Information Technology Conference, Syracuse, NT, pp.121-124, September 1999.
[3] V. Gorodetski, O.Karsaev, A. Karsaev, A. Khabalov, I. Kotenko, L. Popyack, V. Skormin, "Agent-based Model of Computer Network Security System : A Case Study", Proc. of the International Workshop "Mathematical Methods, Models and Architectures for Computer Network Security", LNCS, Vol. 2052, pp.39-50, Springer, 2001.
[4] C.C. Liu, J. Jung, G. Heydt, V. Vittal and A.G. Phadke, "The Strategic Power Infrastructure Defense (SPID) System: A Conceptual Design", IEEE Control Systems Magazine, pp. 40-52, August 2000.
[5] C. Rehtanz, Autonomous Systems and Intelligent Agents in Power System Control and Operation, Springer, New York, 2003.
[6] Rolf Carlson, Sandia SCADA Program: High-Security SCADA LDRD Final Report, Sandia Report, SAND 2002-0729, April 2002.
[7] Thomas Kropp, "System Threats and Vulnerabilities - An EMS and SCADA Security System Overview", IEEE Power and Energy Magazine, pp.46-50, March 2006.
[8] V.M. Igure, S.A. Laughter, R.D. Williams, "Security issues in SCADA networks", Computer & Society, Vol. 25, pp.498-506, 2006.
[9] Avinash Dixit, Susan Skeath, "Games of Strategy", W. W. Norton
[10] Sheldon M. Ross, "Introduction to Probability Models", Academic Press, Inc.

Biography

Hak-Man Kim

Received his B.S., M.S. and Ph.D. degrees in Electrical Engineering from Sungkyunkwan University, Korea in 1991, 1993 and 1998. He was with the Korea Electro-technology Research Institute from Oct. 1996 to Feb. 2008 as a senior researcher. Currently, he is a professor in the Department of Electrical Engineering, Incheon City College, Korea. His research interests include power system engineering, smart power system and agent engineering.
Dr. Kim is a member of KIIEE, KIEE and IEEE.

Dong-Joo Kang

Recieived his B.S. and M.S degrees in Electrical Engineering from Hong-ik University. He has been working for KERI(Korea Electro-technology Research Institute) since 2001. His research interests are on electric power system operation, electricity markets, optimization, operation research, SCADA, cyber security, etc.