Estimation Model of labor Time at the Information Security Audit and Standardization of Audit Work by Probabilistic Risk Assessment

Naoki Satoh, Hiromitsu Kumamoto

Abstract— Based on the factors that are available at the initial phase of the audit task, this paper proposes a labor time estimation method for the information security audit in the form of formula, statistically analyzing the data of the past 20 cases. Initially, audit mode, operation mode, penetration degree, and company size are considered to be the factors that could influence the labor time, and thus the "quantitative analysis I" is conducted with these factors. However the results were not sufficiently positive. As a result, by dividing audit mode into regular and emergency audit and by using company size as the factor, labor time estimation formula has been established by means of the regression analysis. Compared to regular audit, it is found that emergency audit takes more labor time at information security audit. We try to investigate this factor by probabilistic risk assessment.

Keywords— Initiating event, Probabilistic risk assessment, Event tree, Fault tree, Safety measures, Information security

I. Introduction

Recently information security measures have become the important issue that the management should seriously deal with because accidents relating to information security exert great influence on the corporate confidence and thereby on corporate economy. One of the business processes of information security control system is information security audit [1]. Here, the term 'information security audit' means the judgment and advice by the independent information security experts, who scrutinize and appraise whether or not the risk control of an organization is appropriately conducted based on the risk assessment [2] [3]. In order to effectively conduct information security audit, it is necessary to estimate the labor times. However, the estimation has traditionally been dependent on the experiences and instincts of skilled SE and there is no method to estimate labor times quantitatively. Moreover, the accuracy of such estimation by skilled SEs(system engineers) is at 15%-error level at most.

On the other hand, regarding the audit estimation for the development of business software, a number of methods have been propounded such as Function Point, COCOMO, DOTY, PUTNAM, LOC, and so forth. In the estimation of the labor times of information security audit, quantitative analyses based on of many past cases are desirable.

Therefore, by analyzing quantitatively a number of past cases, this paper proposes a method to estimate the labor times of information security audit that can be used at the initial phase of the audit.

II. Information security Audit

A. The Procedure

The procedure of information security audit consists of 4 phases: the planning phase, the implementation phase, the reporting phase, and the improvement phase. This procedure has a cyclical feature, and the total labor times found in this procedure become the factor that is used in the estimation of our information security audit method.

In the planning phase, the plan for document audit and on-the-spot audit is made by extracting necessary audit items according to the purpose of each audit. The specific work of this phase includes grasping what kinds of business the company is doing, identifying where the necessary data exist, determining the range of audit, and so on. Thus, the amount of audit work greatly varies according to the size of the target company and/or the attitude of the target company toward information security.

In the implementation phase, each item is audited under the audit plan. The work is divided into the interview regarding audit items and on-the-spot audit. In the reporting phase, the results of the audit in the implementation phase are documented and reported to the organization that is in charge of scrutiny. This report also includes the evaluation of the information security, the incompatible points, the requirements for improvement, and so on.

In the improvement phase, a plan is made in order to

Naoki Satoh is with Kyoto University, Graduate School of Informatics e-mail: Sato@sys.i.kyoto-u.ac.jp.
Hiromitsu Kumamoto is with Kyoto University, Graduate School of Informatics (e-mail: kumamoto@i.kyoto-u.ac.jp).

improve the audit items that have been judged as incompatible to the audit criteria. The labor times of the information security audit, which is estimated in this paper, is the total number of the labor times in each of these four phases.

B. Influential Factors on Labor Times

In order to make the master plan of information security audit on the basis of labor times, the estimation of the labor times is conducted just before the starting of the planning phase. Since much information cannot be obtained at the starting point of the planning phase, it is necessary to determine the factors that can be considered to influence on the labor times from among the factors that can be obtained at this point.

Such factors as the type of business (manufacturer, service industry, financier, distributors, etc.), the audit form (urgent or regular audit), operation mode (computerized systematic routine or not), penetration degree of information security management, the company size, the location of the target company and so on can be considered to be influential on the labor times we are going to estimate. Among them, four factors can be considered as most influential on the labor times of information security audit: the audit form, operation mode, the penetration degree of information security management, and company size.

(1)The audit form

There are two forms of audit: urgent audit and regular audit. Urgent audit targets the company that has experienced an accident such as the leak of information; therefore, this audit is conducted urgently neglecting the schedule. It is predicted that more labor times will be required in this audit because of the investigation into the accident.

(2)Operation mode

Here, operation mode means whether information security is systematic or not. From the viewpoint of information security audit, companies are divided into three types: company whose security management is systematized, company whose security management is implemented only by documents, and company whose security management is done by both computers and documents. The more systematized the business is, the more efficiently information security audit can be implemented.

(3)The penetration degree of information security management

This degree means to what extent information security is penetrated into management. To be concrete, the more the security control system (such as the establishment of security committee, of security policy, and of security organization) is penetrated, the better the information security management system is. As a result, the labor times of the audit decrease.

(4)The company size

It is likely that the larger the company size is, the number of labor times increases because the number of the audit items and the amount of data to be investigated increase.

III. ACTUAL AUDIT CASES TO BE ANALYZED IN THIS PAPER

In order to calculate the labor times of information security audit, we have collected 20 actual audit cases in the past shown in Table 1. In this paper, the unit of the labor times is man-hour. In this table, the labor times are estimated by system engineers who engaged the audit projects, and its accuracy is indicated by 5 man-hours. Among these 20 cases, no company conducted information security only in the form of document. The penetration degree of information security of the company is subjectively judged as "High" if a security management system is established in the company. Otherwise, judged "Low".

IV. MULTI-VARIABLE ANALYSIS

A. Analysis by Quantitative Analysis Ⅰ

This paper analyzed 20 cases in the past, all of which were equipped with computerized systematic routine. Our hypothesis is that there exists close correlation between the 4 influential factors above (see section Ⅱ.B) and the labor times.

Thus this paper analyzed 17 cases in Table 1 (Company A to Q) with "quantitative analysis I", which can deal with qualitative data and set up a formula to estimate the necessary labor times. Then with the data of the rest 3 cases in Table 1 (Company R to T), the validity of our formula was examined.

In order to analyze by quantitative analysis I, company size, which is a continuous value, is categorized dispersedly as follows:

1) "Very Big": 10,000 employees and above
2) "Big": 5,000 to 9,999 employees
3) "Middle": 1,000 to 4,999 employees
4) "Small": less than 1,000 employees

Table1:Actual Audit Cases

	labor time man_hour	audit form	opetration mode	penetration degree	company size
A	250	urgent	system	low	9219
B	100	urgent	complex	high	220
C	150	urgent	system	low	496
D	200	regular	system	high	9500
E	150	regular	system	low	2100
F	160	regular	system	low	3800
G	140	regular	system	low	3700
H	150	regular	system	low	3500
I	280	regular	system	high	28000
J	115	regular	system	low	200
K	120	regular	system	low	300
L	165	regular	system	low	3500
M	110	regular	complex	low	300
N	170	regular	system	low	5600
O	125	regular	system	low	300
P	160	regular	system	low	5600
Q	200	regular	system	low	12500
R	200	regular	system	low	14000
S	120	regular	system	high	350
T	120	regular	complex	low	530

‘ ‘ Small” companies are likely to have only one business cite, while “Middle” ones are likely to have plural business cites. “Big” companies tend to have many business cites in Japan. And “Very Big” firms usually have more than 10 business cites through out the country and its network system varies from company to company.

In order to determine the formula to estimate the labor times of information security audit, the 4 influential factors on the labor times (see 2.2) are transformed in values dispersedly as follows:

(1) The audit form;
Urgent audit: x11=1,
Regular audit: x12=1
(2) Operation mode;
Full-computerized system: x21=1,
Partial-computerized system: x22=1
(3) Penetration degree;
High penetration: x31=1,
Low penetration: x32=1
(4) Company size;
Very big: x41=1,
Big: x42=1,
Middle: x43=1,
Small: x44=1

Based on the definition above, the formula to estimate the labor times of information security audit of the 17 companies in Table 1 (Company A to Q) is determined as follows:

Labor times =160.6+27.7x11+(-5.94x12)+4.67x21
+(-35.1x22)+(-3x31)+14.5x32
+(-37.1x41)+(-1.2)x42+26.69x43
+74.988x44 (1)

Table2: Actual and the estimated labor times

	estima labor time	actual labor time man_hour	error	audit form	operation mode	penetra degree	company size
A	252	250	2	urgent	system	low	big
B	102	100	2	urgent	complex	high	small
C	146	150	4	urgent	system	low	small
D	198	200	2	regular	system	high	big
E	155	150	5	regular	system	low	middle
F	155	160	5	regular	system	low	middle
G	155	140	15	regular	system	low	middle
H	155	150	5	regular	system	low	middle
I	280	280	0	regular	system	high	very big
J	121	115	6	regular	system	low	small
K	121	120	1	regular	system	low	small
L	155	160	5	regular	system	low	middle
M	108	110	8	regular	complex	low	small
N	155	170	2	regular	system	low	middle
O	121	125	4	regular	system	low	small
P	150	160	10	regular	system	low	big
Q	200	200	0	regular	system	low	very big

penetra.degree: penetration degree
estima labor time: estimated labor time

The comparison between the estimated labor times by the formula (1) and actual labor times is indicated in Table 2. The error level was 2.8%, and the multiple correlation coefficient was 0.91. Therefore, it could be said that the accuracy of the estimation of labor times with the formula (1) is high enough to be used practically. Since it is statistically considered that the nearer the multiple correlation coefficient is to the value 1.0, the accuracy of estimation is high and that the multiple correlation coefficient of a model that can be used practically is more than 0.85, the accuracy of the formula (1) can be considered high enough.

However, as is shown in Table 3, the results of the quantitative analysis I indicate that the partial correlation coefficient of penetration degree is 0.27, which means penetration factor does not influence so much on the labor times. Likewise, the partial correlation coefficient of operation mode is 0.46, which means that operation mode does not influence strongly on the labor times, either.

Table3: Results of labor time estimation by quantitative analysis 1 (4 factors)

item	category	category score	partial correlation coefficient
audit form	urgent	27.7	0.52
	regular	-5.9	
operation mode	system	3.68	0.46
	complex	-27.62	
penetration degree	high	-3.1	0.27
	low	14.5	
company size	very big	74.98	0.83
	big	26.69	
	middle	-1.2	
	small	-37.1	

By neglecting the operation mode factor and the penetration degree factor, we have two influential factors on the estimation of the labor times of information security audit: the audit form (urgent or regular audit) and company size. Moreover, since company size is a quantitative entity, it is possible to seek for the formula to estimate the labor times with regression analysis according to the audit form.

The formula is as follows:

For the case of regular audit,
Labor times = 127.1 + 0.0058*y (2)

For the case of urgent audit,
Labor times = 119.5+0.0142*y (3)

y: company size (number of employees)

The evaluation results by formula (2) and (3) are indicated in Table 4 and 5.

The error levels of formula (2) and formula (3) calculated from the data in Table 4 and 5 are considerably low (6.2% and 10.7% respectively). Likewise, the multiple correlation coefficients of formula (2) and (3) are significantly high (0.97 and 0.95 respectively). This indicates that the accuracies of formula (2) and (3) are high enough to be used practically.

In Table 5, the error of company B’s estimation is big. The reason of this big error is considered as follow: This audit was carried out 7 months after a security accident. In the duration between accident and audit, company B promptly improved

several security management processes. As a result, a security management system could help audit actions and audit labor times was not required than usual urgent audit.

Table4: Estimation Results of regular audit by regression analysis

company ID	estimated labor times	labor times	error	company size
D	183	200	17	9500
E	139	150	11	2100
F	149	160	11	3800
G	149	140	9	3700
H	148	150	2	3500
I	291	280	11	28000
J	128	115	13	200
K	128	120	8	300
L	148	165	17	3500
M	129	110	19	300
N	160	170	10	5600
O	129	125	4	300
P	163	160	3	5600
Q	205	200	5	12500

Table5: Estimation Results of urgent audit by regression analysis

	estimated labor times	labor times man_hour	error	company size
A	240	250	10	9219
B	143	100	43	220
C	146	150	4	496

V. DISCUSSION

The formula (2) in section 4.2 was verified with 3 test data (Company R, S, and T) and the evaluation results are shown in Table 6.

Table6: Evaluation results by test data

	estimated labor times	labor times man_hour	error	company size
R	208	200	8	14000
S	129	120	9	350
T	130	120	10	530

The error level of formula (2) calculated from the data in Table 6 is also considerably low (6.1%), which means the high accuracy of the estimation with formula (2). Taking into consideration the fact that the measuring accuracy of labor times is 5 man-hours, these error levels are highly consistent. Furthermore, since it is generally accepted that the error level of labor time estimation by skilled SEs is roughly 15%, the error level of 6.2% with formula (2) can be considered accurate enough to be used practically.

Compared with formula (2) for regular audit, the constant term of formula (3) is bigger than that of formula (2), and multiplier factor is also larger. This means that if the company size is the same, urgent audit requires more labor times than regular audit. This is in consistency with the fact that more labor times are usually necessary for urgent audit.

On other hand, compared to regular audit, it is found that emergency audit takes more labor time at information security audit. We try to investigate this factor by probabilistic risk assessment [4,5,6].

By probabilistic risk assessment, we try to investigate that emergency audit takes more labor time at information security audit, compared to regular audit, at information security audit. We try to investigate this factor by probabilistic risk assessment [7,8,9].

VI. PRA ANALYSIS

In this section, a sample case is discussed; therefore, in regard to the details of PRA, please refer to the literature and our previous study [10,11].

A. A sample case: Firewall

As indicated in Fig 1, Firewall (F/W) is set in order to protect information asset from illegal access. This is a dual system composed of the main F/W, which usually runs, and the standby F/W, which runs when the main F/W is out of order. The break down of the main F/W triggers an alarm, and the operator, who has caught the alarm, switches to the standby F/W.

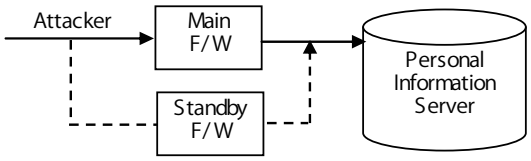

Fig1: Illegal Access and F/W as a Mitigation System

B. Generation of an accident scenario with event trees

As illustrated in Fig 2, in PRA, the scenario of accident occurrence is described with a binary tree called Event Tree, and the point where the two branches diverge each other is called Node. The initiating event is written on the left of the scenario. In this case, the initiating event is “the attempt of an illegal access by the attacker,” and the F/W responses to this initiating event as a mitigation system. In other words, an initiating event can be defined as the event that requires the response of the mitigation system.

To begin with, while the main F/W is working normally, the illegal access can be prevented, which means the mitigation system is working effectively. This is the Scenario 1 in Fig 2.

Next, let us suppose that the main F/W does not work, i.e., it has broken down. In this case, as has been stated in Section

Ⅳ.B an alarm is usually triggered, and the operator detects the abnormality of the main F/W. If the operator is successful in detecting the abnormality, he/she switches to the standby system. The case that the operator succeeded both in detecting the abnormality and in switching to the standby system is Scenario 2 that corresponds to Node 2. Scenario 2 further diverges into another two branches. In the physical system like a nuclear reactor and a chemical plant, the operator has enough time-allowance for switching to the standby system. Therefore, if the operator has successfully detected the breakdown of the main system and switched to the standby system, the accident can be prevented.

However, in the case of information security, it is possible for the attacker to access during the time slot between the break down of the main system and the time when the standby system begins to work. Thus, Scenario 2 further diverges. In Scenario 2.1, illegal access is prevented because both the detection of the abnormality of the main system and the switching to the standby system are successful. In Scenario 2.2, illegal access is not prevented during the time slot between the breakdown and switching, even though both the detection of the abnormality and switching were successful.

As for the length of the blank time slot in the numerical example that will be stated later in Section Ⅳ.D, for the sake of simplicity, it is assumed that it takes 5 minutes to detect the abnormality of the main system and 5 minutes to switch to the standby system; that is, the total length of the blank time slot is 10 minutes. In this example, this time slot length is long enough for the attacker to illegally access because our aim is to explain PRA. Therefore, it goes without saying that depending on the way of access, it can be impossible for the attacker to access.

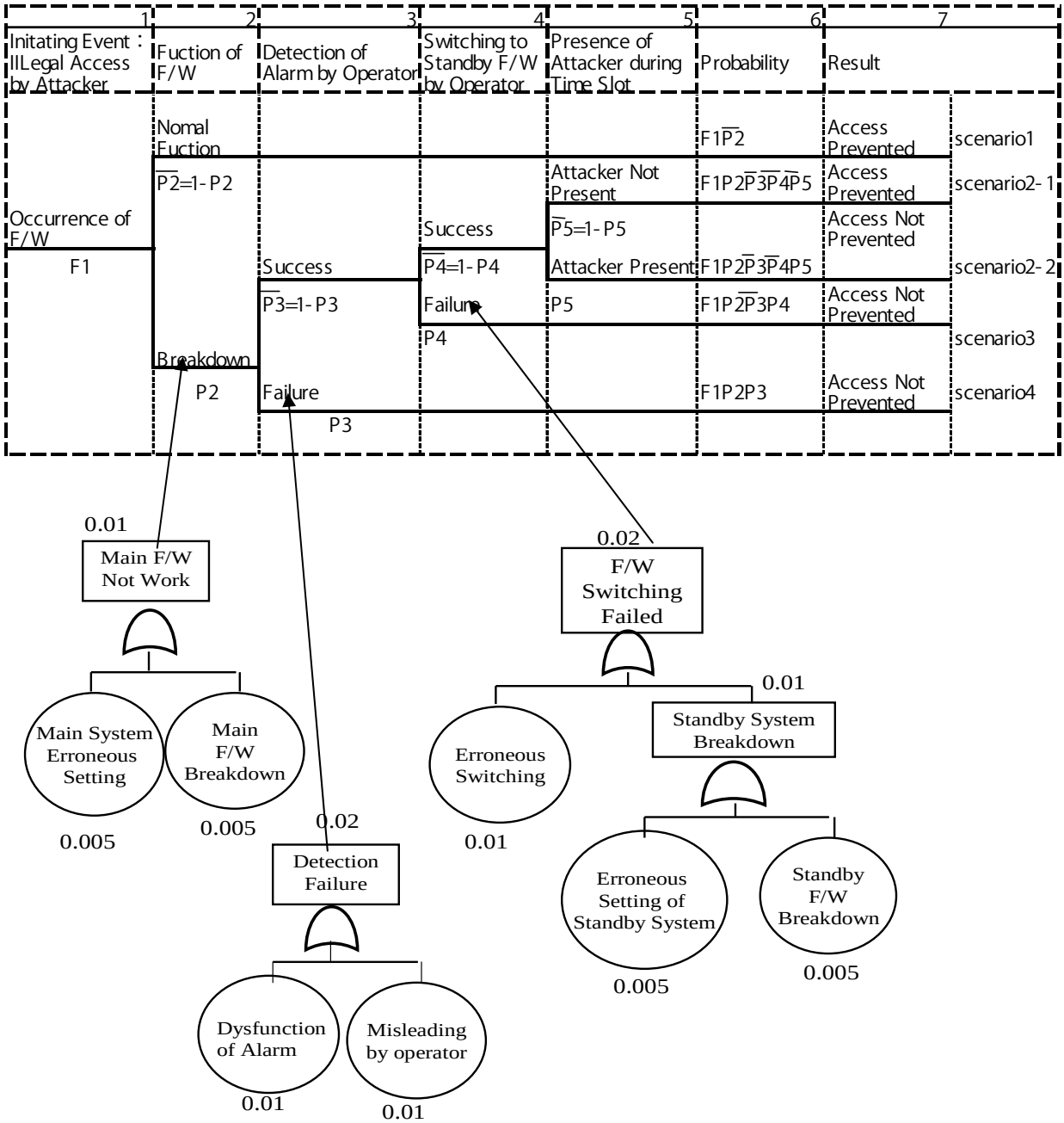

Fig2: Event Tree and Fault Tree of Illegal Access as Initiating Event, F/W example

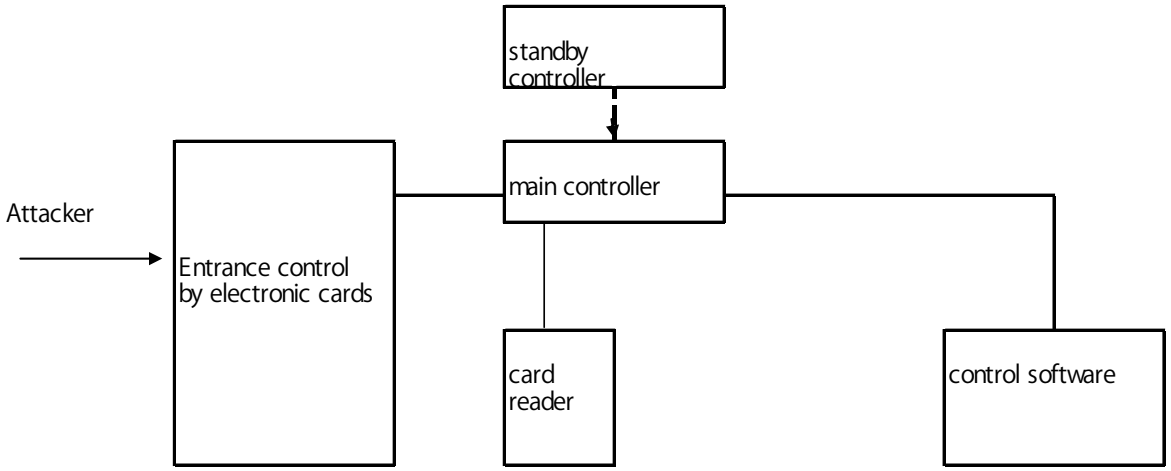

Fig3: Illegal Access and Entrance Control System by Electronic Cards as a Mitigation system

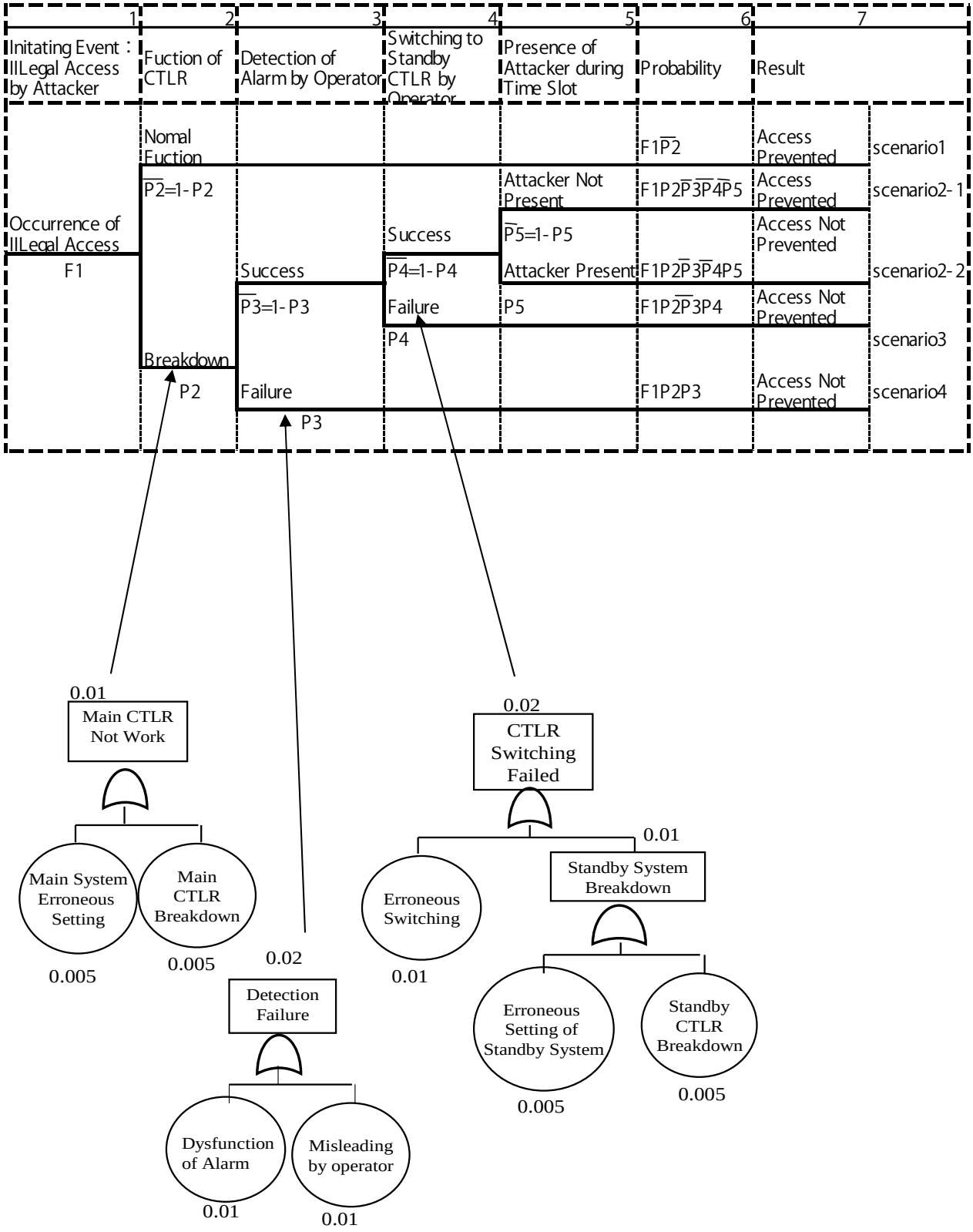

Fig 4 Event Tree and Fault Tree of Illegal Access as Initiating Event, Entrance control

Now let us suppose for the sake of simplicity that the inspection cycle of the dual F/W is one month, that the two F/Ws come back to the mint condition after the inspection, and that the initiating event of "the attempt of illegal access by the

attacker presents during the half of the one-month inspection cycle.

If the initiating event exists during the blank time slot, illegal access is possible. For example, the occurrence frequency of illegal access per month is 1 % in Scenario 2, the possible access frequency per month in Scenario 2.2 is 0.5 %. Needless to say, in Scenario 2.1, because the standby F/W is normally working, illegal access is prevented despite the presence of the initiating event.

In Scenario 3, the detection of the breakdown of the main system was successful but switching to the standby system failed. In this case, the standby F/W does not work and, as a result, illegal access cannot be prevented. From the viewpoint of maintenance, the situation that illegal access cannot be prevented continues until the next routine inspection. Likewise, in Scenario 4, since the detection of the abnormality of the main F/W has failed, illegal access cannot be prevented until the next routine inspection. In Section Ⅳ.D, we will discuss the occurrence frequencies of these scenarios.

C. Analysis of the cause of branching with Fault Tree

The diagram in the lower part of Fig 2 is called Fault Tree that is used for the analysis of the reasons why each Event Tree diverges downwards.

As an example of Fault Tree of the dysfunction of the main F/W, the breakdown of the main F/W itself is a Fault tree on the one hand, which stems from the breakdown of either the hardware or the soft ware, and on the other hand, the mistake in setting the main F/W is also a Fault Tree.

Likewise, as for the cause of the failure of the detection of the breakdown of the main system, the dysfunction of the alarm and the misleading by the operator are the Fault Trees. In addition, as for the cause of the failure of switching to the standby system, erroneous operation and the breakdown of the standby system F/W are the Fault Trees. The latter can be divided into the breakdown of the main F/W itself and the error in setting the main F/W.

The events that are located at the bottom of the Fault Tree are called Basic Events, and in PRA, it is assumed that occurrence frequency and/or occurrence probability can be assigned.

Here, if we assign the numerical values to Basic Events in Fig 2, and if we assume that these events are independent each other, we can approximate the Top Event. For example, let us suppose that the occurrence frequency of the breakdown of the main F/W is 0.0005 times, and that the occurrence frequency of the breakdown of the main F/W that is caused by other reasons than erroneous setting is 0.005 times. Then, it can be approximated that the occurrence frequency of the breakdown of the main F/W is 0.01. Likewise, if it is assumed that the probability of the dysfunction of the alarm under the condition that the main F/W is broken down is 0.01, and that the probability of the erroneous recognition of the alarm by the operator is 0.01, then, it can be approximated that the probability of detection error (so-called Demand Breakdown Probability) is 0.02. Moreover, if it is assumed that the probability of switching failure under the condition that the detection is successful is 0.01, that the probability of the breakdown of the standby F/W caused by the erroneous setting is 0.005, and that the probability of the breakdown of the standby F/W caused by other reasons is 0.005, then it can be approximated that the probability of switching failure after the success of detection is 0.02. *In addition, when the same person set both the main system and the standby system by copying, the dysfunction of the main system means the dysfunction of the standby system, and thereby illegal access cannot be prevented. In this case, the independence of the Basic Events cannot be assumed; therefore, it is necessary to quantify based on the Minimal Cut Set, a failure mode. For example, the pair of the two Basic Events, i.e., the erroneous setting of the main F/W and the dysfunction of the alarm, is a Minimal Cut Set, and is also one of the failure modes of the dual F/W. Therefore, its occurrence frequency can be attained by multiplying the probability or the frequency of the Basic Events. In general, since there exist several Minimal Cut Sets, the scenario is quantified as the total of the occurrence frequency of each Cut Set.*

Finally, the probability varies according to the different cases such as when the same person set the main F/W and the standby F/W individually without copying or when different persons set the main system and the standby system; therefore, it is possible to quantify the safety measures even though it is a relative estimation. Likewise, in the case of alarm detection, the scenario can be assumed that either the operator or the automatic switching worked or not.

D. Analysis with concrete numerical numbers

As is indicated in Fig 2, if it is supposed that the breakdown frequency of the main F/W is 0.01 times per month, the probability of the detection failure is 0.02, and the probability of the switching failure after the successful detection is 0.02, the occurrence frequency under the presence of the initiating event is 0.0096, because $0.01\times0.98\times0.98=0.0096$. If this scenario occurs, since it is assumed that it takes 10 minutes to finish switching, the expected value of the time slot is 0.096 minutes, because$0.0096\times10=0.096$.

Here, in order to exemplify, let us suppose that the real initiating event of the illegal access by the attacker occurs during half of the time slot, then by multiplying 0.096 (the expected value) by 0.5 (the probability of the presence of the initiating event), we can gain 0.048 minutes, which is the time length of illegal access per month in scenario 2.2. In other words, it can be estimated that during 0.048 minutes in a given month, illegal access of scenario 2.2 occurs. In order to reduce this time length, reduction of the time necessary for detection and switching can be considered.

Likewise, in scenario 3, the occurrence

VII. Physical Access Attacker

Consider an entrance control by electronic cards as indicated in Figure3. A duplicated entrance controller permits entrance for personnel with an authorized card. A main controller, an operator, and a standby controller constitute a mitigation system. The event tree is shown in Figure 4. Note that this tree has the same structure as Figure 2 in spite of the fact that the former deals with physical access, while the latter with network access. This indicates that, once an event tree is constructed, a similar version can be applied to other problems of information security.

VIII. Conclusion

In this paper, a model for estimating the labor times for information security audit has been proposed. This method employs the input of the number of employees into the proposed equations at the time of both urgent and regular audit. The results of the analyses of past audit cases indicate that the error levels of this method were 11.4% (urgent audit) and 6.2% (regular audit). It could be concluded that this method has enough practical accuracy, taking into consideration the fact that the error level of the audit by experienced SEs.

As a result, it will be possible to conduct the estimation of information security quantitatively, instead of relying on the traditional estimation that was based on skilled SEs' experience and instinct.

On other hand, compared to regular audit, it is found that emergency audit takes more labor time at information security audit. We try to investigate this factor by probabilistic risk assessment. Following the method of PRA, we have attempted to quantify the risk of information asset by describing a scenario based on the responses of the mitigation systems to the initiating event of each Event Tree and Fault Tree. To be concrete, we supposed a case that an illegal access to the dual F/W, described its scenarios, calculated the occurrence probability of each scenario, and calculated the expected value of the time length of the illegal access.

As a result, it has been quantitatively revealed that to what extent the reduction of the time lengths of switching to the standby system, of the inspection, and of the probability of the failure in detecting dysfunctions and switching exerts influence on the expected value.

IX. Acknowledgement

This paper could not have been completed without various useful advices from my project members and party involved. I would like to express my sincere gratitude to these people.

References

[1] Naoki Satoh and Norihisa Komoda; An Analysis of Influential Factors for the Information Security Audit Labor Time and Regressive Estimation of the Labor Times, WSEAS Trans. on Information Science and Applications, Issue 1, Vol.3, pp.154-161 (2006)

[2] Matsuki Yoshino, Norihisa Komoda, and Michiko Oba: An Analysis of Patterns for Automating Information System Operations, WSEAS Trans. on information science and Application, Issue 11, Vol 5, pp.1618-1627 (2008)

[3] Viewpoint of ISO GMITS and Probabilistic Risk Assessment in information Security, WSEAS Trans on information science and Application, issue 4 , Vol.2, pp237-244(2008)

[4] ASME: Standard for probabilistic risk assessment for nuclear power plant applications, ASME RA-S-2002 (2002)

[5] H. Kumamoto: Modern Reliability Engineering, Corona Inc. (2005)

[6] USNRC: Reactor safety study: An assessment of accident risk in U.S. commercial nuclear power plants. USNRC, WASH1400, NUREG-75/014(1975)

[7] G.E. Apostolakis, J.H. Bickel, S. Kaplan: Probabilistic risk assessment in the nuclear power utility industry, Reliability Engineering and System Safety, vol. 24 no. 2,91-94(1989)

[8] H. Kumamoto, E.J. Henley: Probabilisticrisk assessment and management forengineering and scientists, IEEEPress(1996)

[9] H. Kumamoto: Satisfying safety goals by probabilistic risk assessment, Spinger(2007)

[10] N. Satoh & H. Kumamoto, Enumeration of initiating events of information security accidents, 2007 International Conference Innovation & Management, pp. 119-124(2007)

[11] H. Kumamoto, E. J. Henley: Probabilistic risk assessment and management for engineers and scientists, IEEE Press (1996)