A Labor Time Estimation Model for the Information Security Audit by Quantitative Analysis I and Regression Analysis

NAOKI SATOH,
NS and I System Service Corp
11-17, Fuyuki, Koto, Tokyo, 135-0041
JAPAN

NORIHISA KOMODA
Graduate School of Information Science and Technology
Osaka Universitity
2-1, Yamada-oka, Suita, 565-0871
JAPAN

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 the traditional estimation by skilled system engineers with around 15% error, this proposed formula has 6 to 11% error, which means that this formula has enough practical accuracy.

Key-Words: information security audit, audit plan, labor time estimation, quantitative analysis I, regression analysis, labor times

1 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 [1]. One of the business processes of information security control system is information security audit [2]. 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 [4][5]. 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.

2 Information security audit

2.1 The procedures

The procedure of information security audit consists of 4 phases: the planning phase, the implementation phase, the reporting phase, and the improvement phase[6]. As Fig. 1 shows, 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.

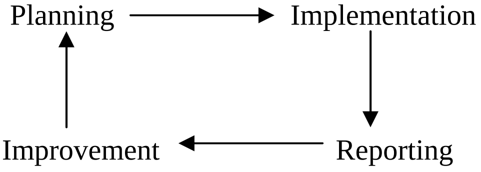

Figure 1: The 4 phases of information security audit

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 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.

2.2 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.

3 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”.

Table 1 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

4 Multi-variable analysis

4.1 Analyses 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 2.2) 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

"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 throughout 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:

$$\begin{aligned}\text{Labor times} =\ &160.6+27.7x11+(-5.94x12)+4.67x21\\ &+(-35.1x22)+(-3x31)+14.5x32\\ &+(-37.1x41)+(-1.2)x42+26.69x43\\ &+74.988x44 \qquad (1)\end{aligned}$$

The comparison between the estimated labor times by the formula (1) and actual labor times is indicated in Table 2.

Table 2: Actual and the estimated labor times

	estimated 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

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.

Table 3: Results of labor time estimation by quantitative analysis I (4 factors)

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

4.2 Regression analysis

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.

Table 4: Estimation Results of regular audit by regression analysis

company ID	estimated labor times	actural 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	129	120	-9	300

L	148	165	17	3500
M	129	110	-19	300
N	160	170	-10	5600
O	129	125	-4	300
P	160	160	0	5600
Q	205	200	-5	12500

Table 5: Estimation Results of urgent audit by regression analysis

company ID	estimated labor times	actual labor times	error	company size
A	251	250	-1	9219
B	122	100	-23	220
C	126	150	23	496

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.

5 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.

Table 6: Evaluation results by test data

	estimated labor times	labor time 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.

Finally, the reason why the operation mode factor and the penetration degree factor, which at first we considered to be influential, have low correlation coefficient, is like this: With regard to the operation mode factor, the more a company is systematized, the less the labor times generally. On the other hand, items to be audited such as computer virus measurement, protection against information leak, encryption and so on are added, as a result of which labor times increase. Therefore, these two aspects offset each other and thus the operation mode factor does not crucially influence on the estimation of the labor times of information security audit. Likewise, with regard to the penetration factor, it is expected that the higher the degree of penetration of security consciousness is, the less the labor times become. However, in reality, management system tends to become complicated together with penetration degree; that is, such a company tends to establish a security committee and improve its security system, as a result of which, in the actual process of audit, it takes longer to conduct the audit and thereby more labor times are required. Thus, the penetration factor can be considered not to be influential.

6 Conclusion

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.

As for the further research, we would like to improve the accuracy of our method by increasing the data for our study.

Acknowledgement:
We would like to express our sincere gratitude to our project members and other people concerning the project, who gave us various kinds of beneficial advice.

References:

[1] The Ministry of Foreign Affairs, *White paper 2004*： *Information and Communications in Japan*, 2004 (in Japanese).

[2] Ministry of Economy, Trade and Industry, *ISMS User's Guide for ISMS Certification Criterion*, 2005 (in Japanese).

[3] Jipdec, *ISMS User's Guide for ISMS Certification Criteria*, 2005 (in Japanese).

[4] Capers Jones, *Estimating Software Costs*, McGraw Hill, 1998.

[5] Shari Lawrence, *Software Engineering Theory and Practice*, Peason Education, 2001.

[6] Ministry of Economy, Trade and Industry, *Information Security Audit Criteria*, 2003 (in Japanese).