Empirical Software Engineering, 10, 495–515, 2005.

An Empirical Approach to Characterizing Risky Software Projects Based on Logistic Regression Analysis

YASUNARI TAKAGI yasunari.takagi@omron.co.jp
Social System Solutions Business Company, OMRON Corporation, Japan

OSAMU MIZUNO o-mizuno@ist.osaka-u.ac.jp
Graduate School of Information Science and Technology, Osaka University, 1-5 Yamadaoka, Suita, Osaka 565-0871, Japan

TOHRU KIKUNO kikuno@ist.osaka-u.ac.jp
Graduate School of Information Science and Technology, Osaka University, 1-5 Yamadaoka, Suita, Osaka 565-0871, Japan

Editor: Michael Deutsch

Abstract. During software development, projects often experience risky situations. If projects fail to detect such risks, they may exhibit confused behavior. In this paper, we propose a new scheme for characterization of the level of confusion exhibited by projects based on an empirical questionnaire. First, we designed a questionnaire from five project viewpoints, requirements, estimates, planning, team organization, and project management activities. Each of these viewpoints was assessed using questions in which experience and knowledge of software risks are determined. Secondly, we classify projects into "confused" and "not confused," using the resulting metrics data. We thirdly analyzed the relationship between responses to the questionnaire and the degree of confusion of the projects using logistic regression analysis and constructing a model to characterize confused projects. The experimental result used actual project data shows that 28 projects out of 32 were characterized correctly. As a result, we concluded that the characterization of confused projects was successful. Furthermore, we applied the constructed model to data from other projects in order to detect risky projects. The result of the application of this concept showed that 7 out of 8 projects were classified correctly. Therefore, we concluded that the proposed scheme is also applicable to the detection of risky projects.

Keywords: Software risk management, questionnaire, risky project, logistic regression.

1. Introduction

Recently, both the functionality and the complexity of software have been increasing as the Information Technology of social systems progresses quickly. In such situations, software development projects are required to produce highly reliable systems within a short period and at low cost. In other words, software development projects have been put in a position of increased risk. Thus, detecting signs of problems related to risk at an early stage of the software project is important. If the detection of a problem is delayed, it becomes more difficult to fix a problem since the effort of coping with a problem increases exponentially with time (Boehm, 1987).

Extensive research has been carried out related to the detection of problems in software development projects. Concerns about risk management are increasing for early detection of such problem signs in software projects. The Software Risk Evaluation method (SRE) is a risk-management technique for software development projects (Williams et al., 1999). In the SRE, the project's risks are identified using the taxonomy table of software risks. The risk taxonomy table is very useful for systematically identifying the risk associated with a project. However, since many risk attributes exist in the identification of risks using the taxonomy table, the extraction of a risk takes time. Therefore, the SRE recommends to carrying out the tailoring of the taxonomy table for each project.

Risks of a software development project are influenced by environmental features, such as the domain, the business style, the culture of the organization, and by human characteristics. In projects with similar environments, an approach to prevent the recurrence of problems by analyzing past problems is usually taken. Such an approach is easily understood by the project members since it is based on problems which actually occur. However, if a problem factor is not arranged appropriately, the number of factors will increase and handling will become more difficult. Consequently, it may causes a disturbance in project management. In order to avoid such a problem, we determined essential problem factors by discussing with project members.

This study presents a statistical analysis of characteristics common to "confused" projects in an organization. The proposed approach has the following features:

Use of a questionnaire: We utilize a simple questionnaire to identify the characteristics of projects.

Empirical evaluation: We empirically validated our constructed model with actual project data.

First, we investigated the problems on various projects using the questionnaire. In developing our questionnaire, we made use of lessons learned from past project experience and related works in the literature (Conrow and Shishido, 1997; Fairley and Rook, 1997; Karolak, 1996; Sisti and Joseph, 1994). From the results of the questionnaire and past project performance, we identified "confused" projects and their problems, and we analyzed such project problems using logistic regression. Logistic regression is a standard classification technique based on a maximum likelihood estimation.

We showed that the proposed model is useful for predicting confused projects in an organization by applying actual project data. In the experiment, we divided the entire data set into two sets according to the time periods in which they were carried out: the data for projects from 1996 to 1997 were used to construct a logistic regression model, while the data for projects in 1998 were applied to achieve an empirical evaluation for the constructed model. We then constructed a regression model using the former data set. In the case of the constructed model, statistical significances and the goodness-of-fit are indicated. The effectiveness of predicting risky projects was subsequently evaluated using the latter data set empirically. The results of this evaluation confirmed the validity of our approach.

The rest of this paper is organized as follows: Section 2 discusses works related to this research. Section 3 defines the notion of confused and risky projects. Section 4 illustrates the outline of our approach to identification and prediction of confused projects. The design of the questionnaire for problem assessment is described in Section 5. The application of the proposed approach to actual project data is shown in Section 6. Section 7 illustrates an empirical evaluation of the effectiveness of the proposed approach. Finally, Section 8 summarizes this paper and discusses future works.

2. Related Works

Risk identification on the software development is a major concern for current software engineering researchers and practitioners. Hence, many lists of risk items have been proposed. For example, Boehm showed the Top-10 risk items for software development (Boehm, 1987). Jones discussed software risk items to be assessed and controlled (Jones, 1993). At the Software Engineering Institute (SEI), a framework to identify and evaluate software risks has been proposed; this framework is called the Software Risk Evaluation Method (SRE) (Williams et al., 1999). Kasser pointed out a list of risk items which cause problems in software development (Kasser and Williams, 1998). In addition, Humphrey related the five most common causes of project failure (Humphrey, 2001).

Although many lists of risk factors have been proposed, the potential risks should be viewed on a case by case basis. In other words, a generic categorization or listing of risk factors needs to be tailored for each development field. We have therefore used a questionnaire to investigate risks and problems in software development within the Social Systems Solutions Business Company (SSBC) at OMRON Corporation. At the same time, several empirical studies investigating software risks by the questionnaire to project managers have also been considered.

Jiang et al. tried to clarify the relationship between software development risks and project effectiveness using the regression model (Jiang and Klein, 2000). The development risks were evaluated by a questionnaire that was mailed to the members of Project Management Institute (PMI). The results of their analysis showed that two risk factors; namely, lack of a team's general expertise and lack of role clarity, were closely related to project effectiveness. However, project effectiveness (as an objective variable) was evaluated by the questionnaire, too. Therefore, less objective results were obtained.

It should be noted that Ropponen et al. investigated the interrelationships among risk factors (Ropponen and Lyytinen, 2000). These authors used principal component analysis and summarized their questionnaire into several large categories. Their work is very useful for summarizing a large questionnaire for risk analysis. The authors also tried to analyze the relationship between risks and environmental factors. That work was mainly aimed at characterizing software risks but not at identifying the risks in software projects.

3. Background and Objective

The main products of SSBC in OMRON consist of embedded software in ticket vending machines, automated teller machines, and in point of sales systems. Such systems for different customers use the same type of hardware and operating systems, but incorporate specific customer requirements, such as customer-specific user interfaces, printed forms, and operation sequences.

The model of development process for these projects is the overlapping waterfall model (Humphrey, 1995).

In order to promote a process improvement initiative, the Software Engineering Process Group (SEPG) was established in 1992. Various Software Process Improvement (SPI) activities were undertaken by the SEPG, such as:

- Collecting and analyzing a project's actual result data such as quality, cost, and duration.
- Monitoring of project status and troubleshooting support.
- Developing process standards and procedures in the organization.
- Delivering training of the organizational standard process.
- Facilitating software process improvement initiatives throughout the organization.

One of the improvement objectives that the SEPG has been striving for is the Project's Planning Accuracy (PPA). PPA is a metric of how well projects performed with respect to their schedule. The PPA is expressed by the ratio of the estimated cost and the planned cost. The PPA reaches 200% when the actual resultant cost reaches twice the planned cost. In general, the PPA becomes higher in confused projects. To improve the PPA, the SEPG has made improvements focused on the project's planning process, such as introducing the Work Breakdown Structure (WBS) and peer review for the project plan (Mizuno et al., 1998). Figure 1 shows the trend of the PPA in the SSBC from 1993 to 2000. As Figure 1 shows, it turns out that the number of projects with about 100% PPA has increased yearly.

However, as shown in Figure 1, the projects with a high PPA (also called "confused"), still occur every year. In order to reduce the number of confused projects, the SEPG has tried to analyze their characteristics. If such characteristics are identified, we can predict confused-prone projects (also called "risky") and we can deal with such projects before they become confused projects. A similar type of project has been dubbed the *death march project* (Yourdon, 1997).

In this paper, a "confused project" is defined as follows:

- a project whose PPA is out of a certain range, and,
- a project falling into an uncontrollable situation during development.

Moreover, we define a "risky project" as a project that may possibly turn into a "confused project." In this paper, we use the term "confused project" to particularly point out the actual status of projects, and the term "risky project" to particularly point out the result of the prediction model.

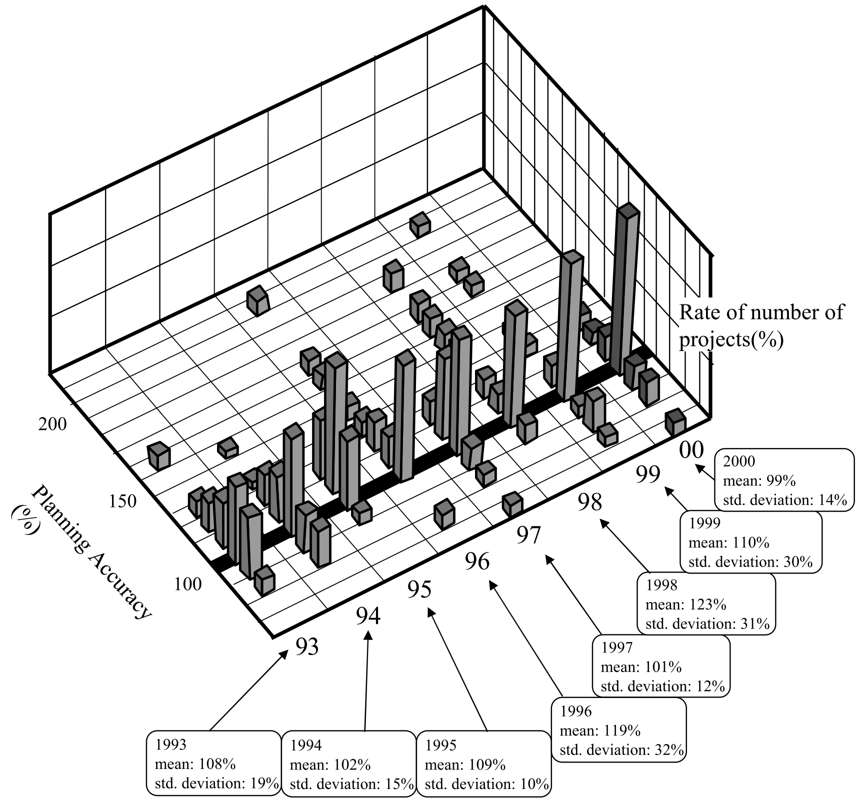


Figure 1. Trend of cost estimation accuracy.

4. Outline of our Approach

Figure 2 shows the outline of our approach for characterizing confused projects. Our approach consists of the following 4 steps. First, in Step 1, we designed a questionnaire to be distributed to project managers and leaders in order to collect the assessment data. Fortunately, in SSBC, actual resulting data and lessons learned from every development project have been stored. The questionnaire consists of five viewpoints each of which is further divided into several risk factors. Next, in Step 2, SEPG distributed the questionnaire to project managers and leaders, and asked them to fill out the questionnaire. After they finished filling out the questionnaire, the SEPG collected them. At the same time, in Step 3, SEPG determined the confused projects from available project data. Here, we assume that the final status of a project becomes either "confused" or "not confused." Finally, in Step 4, from the responses to the questionnaire, we obtained the assessment data, and then constructed a logistic regression model to characterize confused projects.

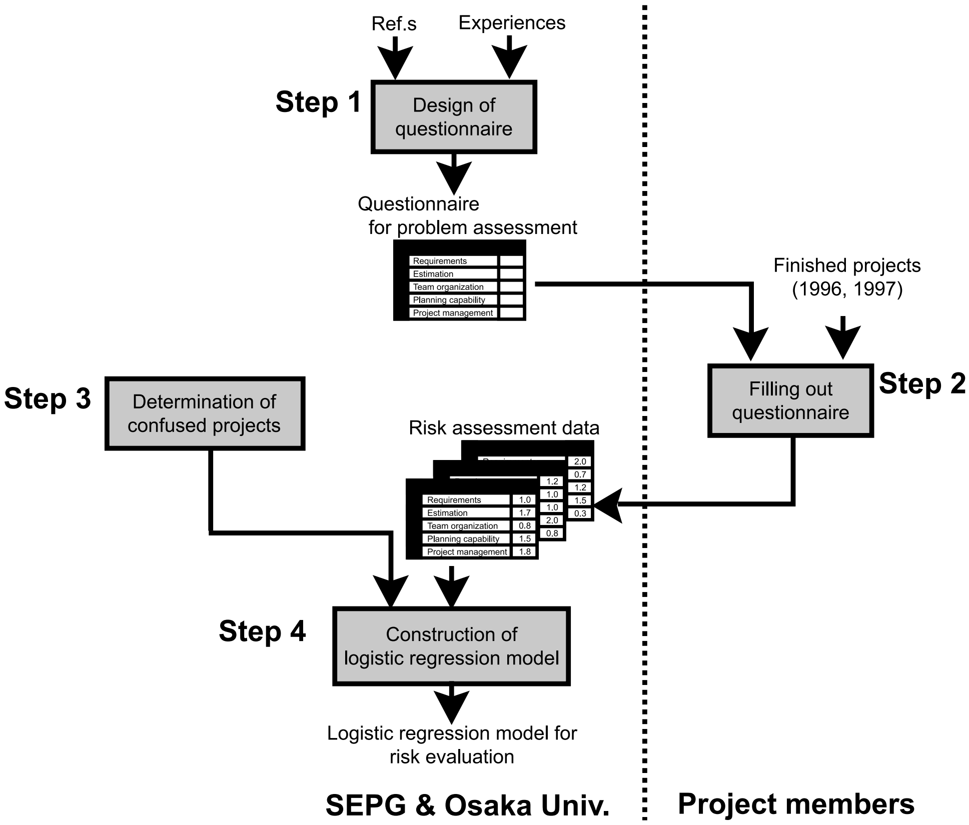


Figure 2. Outline of characterization of confused projects.

The questionnaire consists of five viewpoints: requirements, estimations, planning capability, team organization, and project management activities. (The details of the design will be described in Section 5.) Each sub-item regarding risk factors in the questionnaire must be filled in according to the Likert scale (Fenton and Pfleeger, 1997): “Strongly Agree,” “Agree,” “Neither agree nor disagree,” or “Disagree,”

Logistic regression, a standard classification technique in the experimental sciences, has already been used in software engineering to predict error-prone components (Basili et al., 1996; Briand et al., 1993; Munson and Khoshgoftaar, 1992).

The logistic regression model is based on the following equation:

$$P(Y|x_1,\cdots,x_n) = \frac{e^{b_0 + b_1x_1 + \cdots + b_nx_n}}{1 + e^{b_0 + b_1x_1 + \cdots + b_nx_n}}$$

where $x_1,\cdots,x_n$ are explanatory variables in the model, and Y is a binary dependent variable which represents whether or not a project is confused. P is the conditional probability that Y = 1(i.e., a project is confused) when the values of $x_1,\cdots,x_n$ are determined. We select 22 items in the questionnaire as potential explanatory variables, and estimated the coefficients, or *bi*'s, using the assessment data obtained from the responses to the questionnaire.

In our previous research (Mizuno et al., 2000b), we constructed a regression model with five candidates used for the parameters. In this paper, we try to present a more practical application of this approach. For further analysis, we use detailed factors in the questionnaire to construct a logistic regression model. Using detailed factors is a more efficient way to identify the characteristics of confused projects.

5. Design of the Questionnaire

5.1. Five Viewpoints

In this study, we have investigated various works (Conrow and Shishido, 1997; Fairley and Rook, 1997; Karolak, 1996; Sisti and Joseph, 1994) regarding risk management and the experience of the SSBC. Based on the results of this investigation, we have summarized all key risk factors and classified them into the following five viewpoints: (1) Requirements, (2) Estimations, (3) Team organization, (4) Planning capability, and (5) Project management activities.

The overview of the questionnaire is shown in Figure 3.

5.2. Requirements

The *Requirements* viewpoint includes factors which are related to the understanding and commitment of the requirements among project members. The factors for the requirements viewpoint are distinguished as follows:

(1.1) Ambiguous requirements

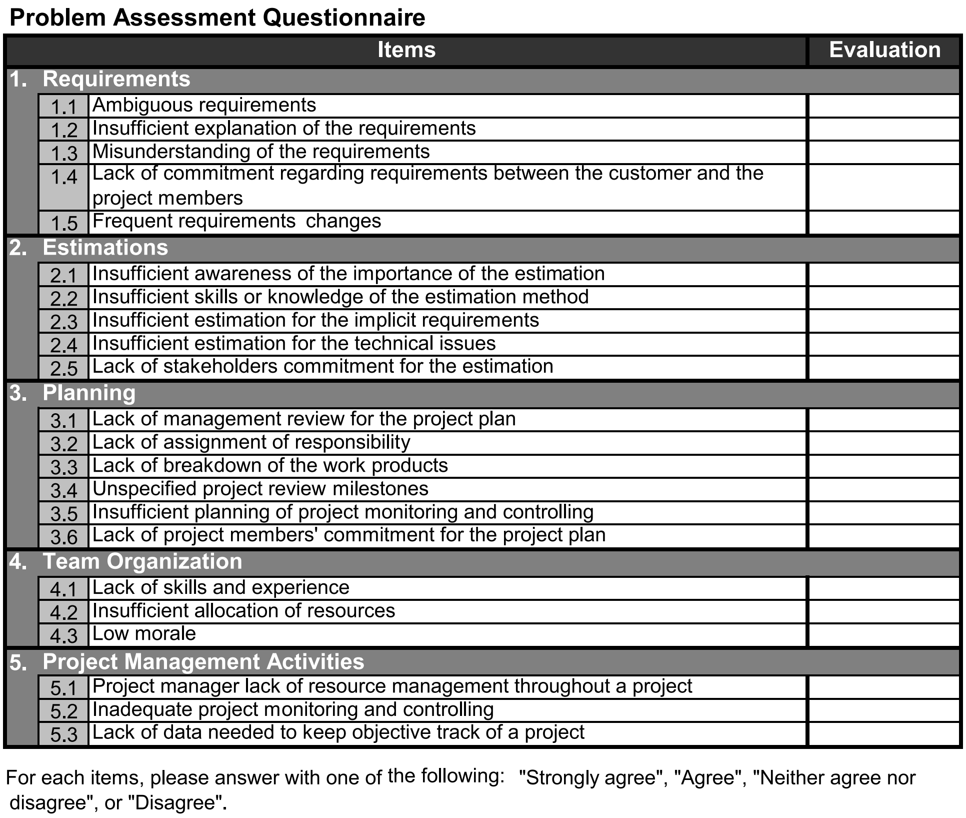
This item checks whether or not the requirements are clear and consistent. It is important for project members to understand what the customer wants in order to achieve clear and consistent results.

(1.2) Insufficient explanation of the requirements

This item checks whether or not customers have a sufficient explanation of the requirements regarding the system and/or software.

(1.3) Misunderstanding of the requirements

This item checks whether or not project members have sufficient skills and/or knowledge to understand the requirements. The developers must have not only sufficient technical skills and/or knowledge for the project, but must also have specific knowledge regarding the customers' domain.



Problem Assessment Questionnaire

	Items	Evaluation
1.	**Requirements**	
1.1	Ambiguous requirements	
1.2	Insufficient explanation of the requirements	
1.3	Misunderstanding of the requirements	
1.4	Lack of commitment regarding requirements between the customer and the project members	
1.5	Frequent requirements changes	
2.	**Estimations**	
2.1	Insufficient awareness of the importance of the estimation	
2.2	Insufficient skills or knowledge of the estimation method	
2.3	Insufficient estimation for the implicit requirements	
2.4	Insufficient estimation for the technical issues	
2.5	Lack of stakeholders commitment for the estimation	
3.	**Planning**	
3.1	Lack of management review for the project plan	
3.2	Lack of assignment of responsibility	
3.3	Lack of breakdown of the work products	
3.4	Unspecified project review milestones	
3.5	Insufficient planning of project monitoring and controlling	
3.6	Lack of project members' commitment for the project plan	
4.	**Team Organization**	
4.1	Lack of skills and experience	
4.2	Insufficient allocation of resources	
4.3	Low morale	
5.	**Project Management Activities**	
5.1	Project manager lack of resource management throughout a project	
5.2	Inadequate project monitoring and controlling	
5.3	Lack of data needed to keep objective track of a project	

For each items, please answer with one of the following: "Strongly agree", "Agree", "Neither agree nor disagree", or "Disagree".

Figure 3. Problem assessment questionnaire.

(1.4) Lack of commitment regarding requirements between customers and project members

This item checks whether or not a commitment is obtained by both project members and customers. In order to confirm a commitment, it is important to have meetings to review requirements with project members and customers.

(1.5) Frequent requirement changes

This item checks whether or not changes in requirements or specifications are appropriately managed and whether or not project members are kept informed.

5.3. Estimations

The *Estimations* viewpoint includes factors related to the estimation itself, the technical methods for carrying out the estimation, and the commitment between project members and customers. The factors for the estimation viewpoint are distinguished as follows:

(2.1) Insufficient awareness of the importance of the estimation

This item checks whether or not project members are aware of the importance of estimations. If they are not aware of the importance of estimation, project members may tend to accept unreasonable requirements.

(2.2) Insufficient skills or knowledge of the estimation method

This item checks whether or not project members have sufficient skills or knowledge of estimation methods. In order to show the rationale of estimates, estimation methods must be utilized effectively.

(2.3) Insufficient estimation for the implicit requirements

This item checks whether or not the implicit requirements are considered and estimated. "Must-be" functions in the business area of the customer, or functions implemented in the previous system tend to be implicit requirements.

(2.4) Insufficient estimation for the technical issues

This item checks whether or not project members have estimated technical issues sufficiently. For example, technical issues include the selection of the programming language and the development environment.

(2.5) Lack of stakeholders' commitment for estimation

This item checks whether or not the commitment between project members and stakeholders has been established. Stakeholders include customers, the sales division, and subcontractors. If a commitment is insufficient and project members yield to political pressure, unrealistic estimations will be produced.

5.4. Planning

The *Planning* viewpoint includes factors related to the planning or scheduling activity and the commitment for the project plan among project members. The factors for the planning viewpoint are distinguished as follows:

(3.1) Lack of management review of the project plan

This item checks whether or not the project manager reviews the project plan. Management review includes reviewing the project plan to check for feasibility, etc.

(3.2) Lack of assignment of responsibility

This item checks whether or not the project has been systematically divided into activities by using the WBS, and whether or not the responsibility for each technical activity has also been clearly specified. The plan of the project should include not only engineering activities but also project management activities.

(3.3) Lack of breakdown of the work products

This item checks whether or not the work products to be produced by development have been correctly specified. The degree of breakdown should be determined for each development project.

(3.4) Unspecified project review milestones

This item checks whether or not sufficient project review milestones have been set up. In the project review, project status, such as progress, quality of work products, etc., is reviewed.

(3.5) Insufficient planning of project monitoring and controlling

This item checks whether or not the plan to monitor and control project activities, issues, risks, and work products is specified correctly.

(3.6) Lack of project members' commitment for the project plan

This item checks whether or not the plan has been reviewed by all of project members. All engineers engaged in the project must recognize the project plan and understand the concrete goals of the project.

5.5. Team Organization

The *Team organization* viewpoint includes factors related to the staffing of the projects, the fundamental skills and experience, and morale of project members. The factors for the team organization viewpoint are as follows:

(4.1) Lack of skills and experience

This item checks whether or not project members have sufficient skills and experience to do their tasks.

(4.2) Insufficient allocation of resources

This item checks whether or not the resources are well allocated.

(4.3) Low morale

This item checks whether or not the morale of project members is low.

5.6. Project Management Activities

The *Project management activities* viewpoint includes factors related to project management activities. The factors which distinguish the project management activities viewpoint are as follows:

(5.1) Project manager lack of resource management throughout a project

This item checks whether or not project members are actually working on the assigned project. Project managers should act as a firewall so that project members can devote themselves to their tasks.

(5.2) Inadequate project monitoring and controlling

This item checks whether or not progress monitoring is adequately done, whether or not progress reporting is actually done, and whether or not corrective action is adequately taken.

(5.3) Lack of data needed to keep objective track of a project

This item checks whether or not project managers are able to objectively keep track of a project on the basis of the software metrics collected during development. If such data are not available, then it is difficult to recognize the project status correctly to make management decisions appropriately.

5.7. Comparison Between Past Studies

As mentioned previously, the proposed questionnaire is based mainly on the experiences of project members and the past studies of the software project risks.

Here, we compare the proposed questionnaire with the lists of risk factors in the literature (Kasser and Williams, 1998; Boehm, 1987; Humphrey, 2001). For the comparison, we use two recent risk lists, one from Humphrey (2001) and the other from Kasser (1998). Humphrey's list is shown in his book and it is widely known in this field. On the other hand, Kasser's risk list is based on interviews with developers in DoD, and thus includes empirical viewpoints. In this sense, our questionnaire is similar to Kasser's list.

Figure 4 shows the correspondence between our list and Humphrey's and the one between our list and Kasser's. We can see that our proposed questionnaire covers a large number of risk factors included in both lists. For example, factors (2.3) and (2.4)

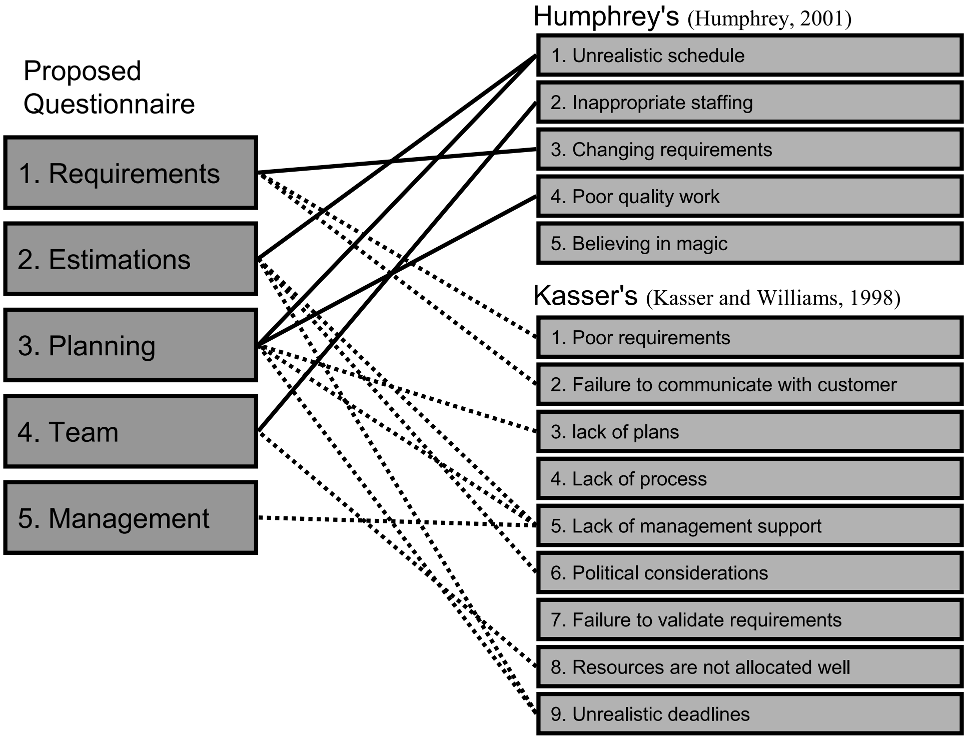

Humphrey's (Humphrey, 2001)
1. Unrealistic schedule
2. Inappropriate staffing
3. Changing requirements
4. Poor quality work
5. Believing in magic
Proposed Questionnaire
1. Requirements
2. Estimations
3. Planning
4. Team
5. Management
Kasser's (Kasser and Williams, 1998)
1. Poor requirements
2. Failure to communicate with customer
3. lack of plans
4. Lack of process
5. Lack of management support
6. Political considerations
7. Failure to validate requirements
8. Resources are not allocated well
9. Unrealistic deadlines

Figure 4. Comparison between the proposed questionnaire and past studies.

correspond to both the "1. Unrealistic schedule" of Humphrey's list and the "9. Unrealistic deadlines" of Kasser's list. As another example, factor (4.2) corresponds to both the "2. Inappropriate staffing" from Humphrey's and the "8. Resources are not allocated well" from Kasser's list.

6. Construction of the Statistical Model

In this section, we apply the proposed questionnaire in Section 5 to sample projects and successively apply logistic regression analysis to the assessment data obtained from the questionnaire.

6.1. Determination of Confused Projects

First, we chose 32 projects, which were part of the projects performed from 1996 to 1997 by the SSBC. Since all of these projects completed their development, the SEPG had already identified the confused projects according to the decision process mentioned in

Section 4. As a result, 10 projects out of 32 were classified as confused projects. Thus, the column "Actual Result" in Table 1 shows the actual result of the classification.

6.2. Filling Out the Questionnaire

The SEPG distributed the questionnaires designed in Section 5 to the project managers or the project leaders of 32 target projects, and explained the details of the questionnaire and the purpose of the trial. The responses to the questionnaire were collected by the SEPG after one month.

In order to elucidate the assessment data from the responses, we assigned points 3, 2, and 0 to "Strongly Agree," "Agree," and "Disagree" for each evaluation (in Figure 3), respectively. We assigned 1 point to the evaluation, "Neither agree nor disagree," because it was assumed that project managers encountered problems that were not quite serious enough to warrant the selection of "Disagree."

6.3. Construction of the Logistic Regression Model

Next, we considered all of the 22 factors in the questionnaire, each of which corresponds to 22 questions in the questionnaire, as candidates of parameters in a logistic regression model.

From the assessment data shown in Table 1, we constructed a regression model. Four parameters were adopted in the regression model using the stepwise method:

2.3 Insufficient estimation for the implicit requirements (x_1)

2.5 Lack of stakeholder's commitment for estimation (x_2)

3.3 Lack of breakdown of the work products (x_3)

3.5 Insufficient planning of project monitoring and control (x_4)

The coefficients for these parameters(that is, b_1, b_2, b_3, and b_4, respectively) were determined as shown in Table 2. The p-value in Table 2 represents the statistical significance. It represents the probability that the coefficient is different from zero by chance. A significance threshold of 0.20 has often been used to select parameters in the stepwise regression.

Furthermore, the correlation coefficients among these 4 parameters are not very high, as shown in Table 3. We can say that these 4 parameters are statistically independent, and it is valid to include these 4 parameters in the model.

"Odds ratio" in Table 2 represents the ratio between the conditional probability of being risky and the conditional probability of being not risky when the value of the explanatory variable increases by one unit. Intuitively speaking, this ratio represents the impact of the explanatory variable in making projects risky.

Table 1. Projects in ’96 and ’97 used for our experiment.

	1: Requirements					2: Estimations					3: Planning						4: Organization			5: Management			
Projects in ’96 & ’97	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	6	1	2	3	1	2	3	Results
PJ1	0	0	0	0	0	2	3	3	2	0	2	0	0	0	0	0	2	1	0	0	0	0	Not confused
PJ2	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	Not confused
PJ3	0	0	0	0	3	0	0	2	3	0	0	0	0	0	2	0	0	0	0	0	0	0	Not confused
PJ4	3	3	2	2	3	0	0	2	2	0	2	2	0	0	0	1	2	0	0	0	0	0	Not confused
PJ5	0	0	0	0	2	0	0	0	0	0	0	2	0	2	2	0	0	0	0	0	2	0	Not confused
PJ6	0	3	2	0	0	2	2	2	0	2	0	2	0	0	0	0	0	0	0	0	0	2	Not confused
PJ7	0	0	2	3	2	0	0	0	0	0	0	2	0	3	0	0	0	0	0	0	0	0	Not confused
PJ8	0	2	3	3	0	1	0	2	0	0	2	2	0	0	2	2	0	0	1	3	0	0	Not confused
PJ9	0	2	0	2	3	0	0	0	0	0	2	2	0	2	2	0	0	0	0	0	0	2	Not confused
PJ10	0	0	0	0	2	0	2	2	0	0	0	2	0	0	2	0	0	0	0	2	0	0	Not confused
PJ11	0	3	3	2	0	0	0	3	3	0	0	0	0	0	0	0	0	0	2	0	0	0	Not confused
PJ12	0	2	2	2	0	0	2	0	0	0	0	2	0	2	0	0	0	0	0	0	0	2	Not confused
PJ13	0	2	0	2	0	0	0	0	0	0	2	3	3	0	2	2	2	2	0	2	2	1	Not confused
PJ14	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Not confused
PJ15	0	2	2	2	2	0	2	2	0	0	0	0	0	0	0	0	3	2	0	3	0	0	Not confused
PJ16	0	0	0	0	2	0	2	0	2	3	3	2	0	2	3	2	3	2	0	2	2	2	Not confused
PJ17	0	0	0	0	0	0	2	0	0	0	2	2	2	3	2	2	0	0	0	2	2	0	Not confused
PJ18	0	0	0	0	1	0	0	0	0	0	0	2	2	0	0	0	0	0	0	0	0	0	Not confused
PJ19	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Not confused
PJ20	0	2	3	2	3	0	0	0	0	0	3	0	0	0	3	0	2	0	0	0	3	3	Not confused
PJ21	0	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Not confused
PJ22	3	2	3	3	2	2	1	3	2	1	0	2	2	2	0	1	3	1	2	2	2	0	Not confused
PJ23	2	2	0	2	3	0	0	2	3	0	2	0	2	2	3	2	0	0	0	0	2	3	Confused
PJ24	2	2	3	3	3	2	2	3	2	3	3	3	3	2	3	2	3	3	0	2	2	2	Confused
PJ25	3	2	0	0	3	0	0	0	0	0	3	0	0	3	3	0	0	0	0	0	0	0	Confused
PJ26	0	2	3	2	2	3	0	2	2	1	0	2	0	0	2	2	0	2	2	2	0	2	Confused
PJ27	0	2	2	2	2	0	3	2	3	3	0	2	2	0	0	2	2	2	0	0	0	0	Confused
PJ28	2	3	3	2	2	0	0	3	3	2	3	0	3	0	2	3	2	0	2	0	2	2	Confused
PJ29	3	2	3	2	0	3	2	2	2	0	0	2	2	2	3	0	2	0	2	0	3	3	Confused
PJ30	2	2	3	3	2	0	0	2	0	2	2	2	2	2	2	0	3	0	2	0	2	0	Confused
PJ31	0	0	0	0	0	0	0	0	0	0	3	3	3	3	3	3	3	3	0	0	3	3	Confused
PJ32	2	3	3	3	2	2	3	3	3	3	3	3	3	2	3	3	3	3	2	3	3	0	Confused

Table 2. Coefficients in the logistic regression model.

		Coefficient	Odds ratio	p-value
	Intercept	−8.834 (b$_0$)	0.000	–
x_1	2.3 Insufficient estimation for the implicit requirements	1.577 (b$_1$)	4.840	0.124
x_2	2.5 Lack of stakeholder's commitment for estimation	0.964 (b$_2$)	2.622	0.126
x_3	3.3 Lack of breakdown of the work products	1.228 (b$_3$)	3.414	0.074
x_4	3.5 Insufficient planning of project monitoring and controlling	2.222 (b$_4$)	9.226	0.054

The statistical significance of this formula is confirmed by 0.01 level.

We then investigated the significance of the whole model. First, the hypothesis H_0 : $b_1 = b_2 = b_3 = b_4 = 0$ was tested by the likelihood ratio test. By the likelihood ratio test, it was shown that the p-value of the model was less than 0.0001. This model was thus shown to be statistically significant.

As a result, we obtained the following expression in order to calculate the conditional probability for a given project:

$$P(Y|x_1, x_2, x_3, x_4) = \frac{e^{-8.834 + 1.577x_1 + 0.964x_2 + 1.228x_3 + 2.222x_4}}{1 + e^{-8.834 + 1.577x_1 + 0.964x_2 + 1.228x_3 + 2.222x_4}}$$

Table 4 shows the calculated conditional probabilities for projects in 1996 and 1997. Table 4 was divided into two groups: the Confused and Not confused groups. Here, we assume that the classification threshold is 0.5. Table 5 summarizes the prediction by the constructed model. It was shown that 28 out of 32 projects were predicted correctly. Most importantly, 9 out of 10 actually confused projects(they were seen to be "Confused" in Table 4) were predicted correctly.

7. Empirical Evaluation

In order to evaluate the effectiveness of the proposed approach, we performed an empirical experiment by applying the proposed approach to the assessment data set of the 1998's projects. The evaluation process is depicted in Figure 5.

Table 3. Correlation coefficients among parameters in the model.

		2.3	2.5	3.3	3.5
x_1	2.3 Insufficient estimation for the implicit requirements				
x_2	2.5 Lack of stakeholder's commitment for estimation	0.42			
x_3	3.3 Lack of breakdown of the work products	0.27	0.43		
x_4	3.5 Insufficient planning of project monitoring and controlling	0.02	0.22	0.39	

Table 4. Calculated conditional probabilities for projects in 1996 and 1997.

Projects in	2: Estimations		3: Planning				
'96 & '97	2.3 (x_1)	2.5 (x_2)	3.3 (x_3)	3.5 (x_4)	$P(Y/x_1,\ldots,x_4)$	Predicted	Actual result
PJ1	3	0	0	0	0.0163	Not risky	Not confused
PJ2	0	0	0	0	0.0001	Not risky	Not confused
PJ3	2	0	0	2	0.2251	Not risky	Not confused
PJ4	2	0	0	0	0.0034	Not risky	Not confused
PJ5	0	0	0	2	0.0122	Not risky	Not confused
PJ6	2	2	0	0	0.0229	Not risky	Not confused
PJ7	0	0	0	0	0.0001	Not risky	Not confused
PJ8	2	0	0	2	0.2251	Not risky	Not confused
PJ9	0	0	0	2	0.0122	Not risky	Not confused
PJ10	2	0	0	2	0.2251	Not risky	Not confused
PJ11	3	0	0	0	0.0163	Not risky	Not confused
PJ12	0	0	0	0	0.0001	Not risky	Not confused
PJ13	0	0	3	2	0.3305	Not risky	Not confused
PJ14	0	0	0	0	0.0001	Not risky	Not confused
PJ15	2	0	0	0	0.0034	Not risky	Not confused
PJ16	0	3	0	3	0.6735	Risky	Not confused
PJ17	0	0	2	2	0.1263	Not risky	Not confused
PJ18	0	0	2	0	0.0017	Not risky	Not confused
PJ19	0	0	0	0	0.0001	Not risky	Not confused
PJ20	0	0	0	3	0.1027	Not risky	Not confused
PJ21	0	0	0	0	0.0001	Not risky	Not confused
PJ22	3	1	2	0	0.3356	Not risky	Not confused
PJ23	2	0	2	3	0.9690	Risky	Confused
PJ24	3	3	3	3	0.9999	Risky	Confused
PJ25	0	0	0	3	0.1027	Not risky	Confused
PJ26	2	1	0	2	0.4324	Not risky	Confused
PJ27	2	3	2	0	0.4178	Not risky	Confused
PJ28	3	2	3	2	0.9974	Risky	Confused
PJ29	2	0	2	3	0.9690	Risky	Confused
PJ30	2	2	2	2	0.9588	Risky	Confused
PJ31	0	0	3	3	0.8199	Risky	Confused
PJ32	3	3	3	3	0.9999	Risky	Confused

7.1. Filling out the Questionnaire

Here, the questionnaires were again distributed to managers of other projects.

7.2. Prediction Using Logistic Regression Model

We then applied the logistic regression model constructed in Section 6 to the assessment data set for the projects in 1998, and calculated the conditional probabilities of being risky. The result of this experiment is shown in Table 6.

Table 5. Actual results vs. prediction for projects in 1996 and 1997.

	Predicted	
Actual result	Not risky	Risky
Not confused	19	3
Confused	1	9

In Table 6, $P(Y/x_1,\cdots,x_4)$ shows the calculated conditional probability of being risky. We again assume that the threshold of probability of a project being risky is 0.5. We can see that the conditional probabilities(that is, $P(Y/x_1,\cdots,x_4)$'s) of $PJ39$ and $PJ40$ are higher than 0.5. That is, these two projects were predicted to be risky projects. The final status

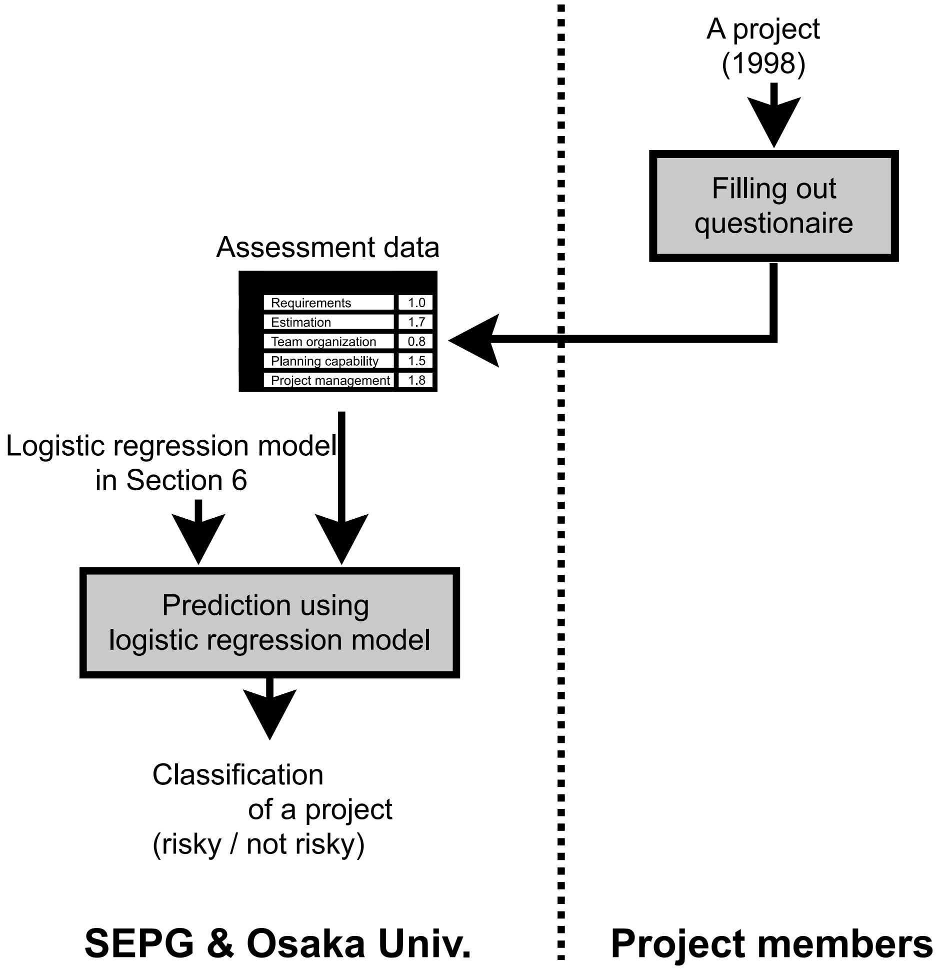


Figure 5. Outline of evaluation process.

Table 6. Calculated conditional probabilities for projects in 1998

	2: Estimations		3: Planning				
Projects in '98	2.3 (x_1)	2.5 (x_2)	3.3 (x_3)	3.5 (x_4)	$P(Y/x_1,\cdots,x_4)$*	Predicted	Actual Result
PJ33	0	0	0	0	0.0001	Not risky	Not confused
PJ34	0	0	0	0	0.0001	Not risky	Not confused
PJ35	0	2	0	0	0.0010	Not risky	Not confused
PJ36	0	0	2	0	0.0017	Not risky	Not confused
PJ37	0	0	2	0	0.0017	Not risky	Not confused
PJ38	2	3	2	0	0.4178	Not risky	Confused
PJ39	3	1	3	0	0.6330	Risky	Confused
PJ40	3	0	3	3	0.9981	Risky	Confused

* $P(Y/x_1,\cdots,x_4)$ shows the conditional probability of being risky project.

of these projects is also shown in Table 6 as the column "Actual Result." In Table 7, by comparing the probabilities with the final status, we can see that 7 out of 8 projects were predicted correctly.

7.3. Discussion

Our objective for this study was to propose an approach to the identification of risky projects (that is, projects prone to be confused) using the questionnaire and the logistic regression model. We discuss this objective here from the following two viewpoints:

(1) *Validity of the model.* The following reasons for the prediction rate of the statistical model having been high in Sections 6.3 and 7.2 are considered:

- The business domain has not changed during 1996 to 1998.
- No big changes in platforms, such as hardware and development environments occurred.
- All projects involved customized development of embedded software.

In other words, the similarity of the environment in which the project was placed is high. Because of this high similarity of environment, we believe that the affinity to the statistical approach is high, and therefore, the prediction rate of applying the model also

Table 7. Actual results vs. prediction for projects in 1998.

	Predicted	
Actual result	Not risky	Risky
Not confused	5	0
Confused	1	2

became high. Also, when the similarity of environment is high, project characteristics can be extracted by the statistical technique. Moreover, the assessment by this statistical approach can narrow down 22 factors to 4 factors. Although many factors can put a project at risk, we are able to show that practical narrowing down can be performed on many risk items in a project or an organization by this approach.

(2) *Extracted characteristics*. In this paper, we determined that 4 specific factors are deeply related to confused projects. We tried to analyze this result in more detail. All of these 4 factors are related to either "estimation" or "planning" viewpoints. This implies a simple principle: most project risk can be avoided if the project plan and the estimates are constructed carefully. As mentioned before, process improvement activities have been performed at the SSBC. One of the most emphasized activities has been adherence to the construction of the project plan (Mizuno et al., 1998; Mizuno et al., 2000a). The results of this experiment supports the validity of such process improvement activities.

Now, we consider "2.3 Insufficient estimation for the implicit requirements," extracted as one of the 4 factors. The projects carried out a customized design of the embedded software. For embedded software, "implicit functions," such as automatic error correction, are more important than "explicit functions." Typically, implicit functions are not specified, but are served as knowledge that is peculiar to people, equipments, domain, etc. The fact that factor 2.3 was extracted shows the above-mentioned features of the embedded domain.

8. Conclusion

In this research, we performed statistical analyses of the problems generated in software development projects and the relationship of project results using the stepwise logistic regression model. Consequently, we showed that a simple prediction model can be constructed for project results (that is, risky or not risky) using a binary classification scheme. Furthermore, an experiment was conducted, which verified the validity of our predictive model. Since data used in this study describe the final status of projects and the responses to the questionnaire about software risk, these data can be easily collected in actual development projects. The extracted factors can be used as helpful information for risk identification at the beginning of a project. In this research, the problem factors were narrowed down to a project management perspective. Future work will explore other aspects of this problem domain such as product engineering and customer relationship management.

Acknowledgments

The authors would like to thank the late Dr. Keishi Sakamoto for his support to our entire process of improvement activities. The authors also would like to thank Associate Prof. Shinji Kusumoto of Osaka University for his discussions and advice regarding the analysis in this paper. The authors would like to express great thanks to the project members of the 40 projects who willingly cooperated with our requests. Finally, the authors would like to thank the anonymous reviewer who gave us invaluable suggestions.

References

Basili, V. R., Briand, L. C., and Melo, W. L. 1996. A validation of object oriented metrics as quality indicators. *IEEE Transactions on Software Engineering* 22(10): 751–761.

Boehm, B. W. 1987. Industrial software metrics top 10 list. *IEEE Software* 4(5): 84–85.

Briand, L. C., Basili, V. R., and Hetmanski, C. 1993. Developing interpretable models with optimized set reduction for identifying high risk software components. *IEEE Transactions on Software Engineering* 19(11): 1028–1044.

Conrow, E. H., and Shishido, P. S. 1997. Implementing risk management on software intensive projects. *IEEE Software* 14(3): 83–89.

Fairley, R., and Rook, P. 1997. Risk management for software development. In *Software Engineering*. Los Alamitos, CA: IEEE CS Press, pp. 387–400.

Fenton, N. E., and Pfleeger, S. L. 1997. *Software Metrics : A Rigorous & Practical Approach*, PWS Publishing.

Humphrey, W. S. 1995. *A Discipline for Software Engineering*. Maryland: Addison Wesley.

Humphrey, W. S. 2001. *Winning with Software: An Executive Strategy*. Maryland: Addison-Wesley.

Jiang, J., and Klein, G. 2000. Software development risks to project effectiveness. *Journal of Systems and Software* 52: 3–10.

Jones, C. 1993. *Assessment and Control of Software Risks*. Englewood Cliffs, NJ: Prentice Hall, Inc.

Karolak, D. W. 1996. *Software Engineering Risk Management*. California: IEEE CS Press.

Kasser, J., and Williams, V. R. 1998. What do you mean you can't tell me if my project is in trouble? *DoD Software Tech News* 2(2). http://www.dacs.dtic.mil/awareness/newsletter/technews2-2/trouble.html

Mizuno, O., Kikuno, T., Inagaki, K., Takagi, Y., and Sakamoto, K. 1998. Analyzing effects of cost estimation accuracy on quality and productivity. In *Proceedings of 20th International Conference on Software Engineering*. pp. 410–419.

Mizuno, O., Kikuno, T., Inagaki, K., Takagi, Y., and Sakamoto, K. 2000a. Statistical analysis of deviation of actual cost from estimated cost using actual project data. *Information and Software Technology* 42: 465–473.

Mizuno, O., Kikuno, T., Takagi, Y., and Sakamoto, K. 2000b. 'Characterization of risky projects based on project managers' evaluation. In *Proceedings of 22nd International Conference on Software Engineering*. pp. 387–395.

Munson, J., and Khoshgoftaar, T. 1992. The detection of faultprone programs. *IEEE Transactions on Software Engineering* 18(5): 423–433.

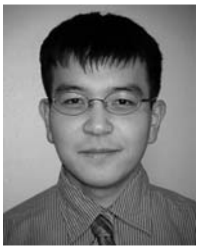
Ropponen, J., and Lyytinen, K. 2000. Components of software development risk: How to address them? A project manager survey. *IEEE Transactions on Software Engineering* 26(2): 98–112.

Sisti, F. J., and Joseph, S. 1994. Software risk evaluation method version 1.0. Technical Report CMU/SEI94TR19, Software Engineering Institute.

Williams, R. C., Pandelios, G. J., and Behrens, S. G. 1999. Software risk evaluation (SRE) Method Description (Version 2.0). Technical Report CMU/SEI99TR029, Software Engineering Institute.

Yourdon, E. 1997. *Death March: The Complete Software Developer's Guide to Surviving 'Mission Impossible' Projects*. Englewood Cliffs, NJ: Prentice-Hall Computer Books.

Yasunari Takagi received B.E. degree in information and computer science from Nagoya Institute of Technology in 1985. He has been working for OMRON Corporation. He is currently a Ph.D. course student in Graduate School of Information Science and Technology, Osaka University. His current research interests include software process improvement activities in industries and software quality assurance of embedded software systems.



Osamu Mizuno received M.E. and Ph.D. degrees from Osaka University in 1998 and 2001, respectively. He is currently an Assistant Professor in Graduate School of Information Science and Technology at Osaka University. His research interests include methodologies of software process improvement, risk evaluation, and prediction in software development. He is a member of IEEE.

Tohru Kikuno received M.Sc. and Ph.D. degrees from Osaka University in 1972 and 1975, respectively. He joined Hiroshima University from 1975 to 1987. Since 1990, he has been a Professor of the Department of Information and Computer Sciences at Osaka University. Since 2002, he has been a Professor of Graduate School of Information Science and Technology at Osaka University. He also holds a Director of Osaka University Nakanoshima Center from 2004. His research interests include the analysis and design of fault-tolerant systems, the quantitative evaluation of software development processes, and the design of procedures for testing communication protocols.

He is a senior member of IEEE, a member of ACM, IEICE (the Institute of Electronics, Information and Communication Engineers), and a fellow of IPSJ (Information Processing Society of Japan).