Characterization of Risky Projects based on Project Managers' Evaluation

Osamu Mizuno
Tohru Kikuno
Osaka University, JAPAN

Yasunari Takagi
Keishi Sakamoto
OMRON Corporation, JAPAN

1

June 9, 2000

ICSE2000 · Limerick · Ireland

Outline

- **Background and objective**

- Definition of "risky" project

- Our approach

- Design of a questionnaire (Step 1)

- Construction of a statistical model (Step 3)

- Risk prediction (Step 5)

Background

- In the company, as a result of process improvement for 8 years, the number of "confused" projects has decreased.
  - However, several projects fall into dangerous state in their development (we call them "risky projects"), and a few of them rarely cause confusion.

- We recognized such confusion by quantitative measure after the project finishes.

- However, prediction of the risky projects should be carried out at the early stage of the project.

Objective

- Our objective is as follows:

   (1) Identify characteristics of a risky project based on the questionnaire and logistic regression model.

   (2) Propose a framework to predict the risky project.

   (3) Apply the proposed model to actual projects.

Outline

- Background and objective

- **Definition of "risky" project**

- Our approach

- Design of a questionnaire (Step 1)

- Construction of a statistical model (Step 3)

- Risk prediction (Step 5)

Risky Project – What's risky?

- "Risky project" is a confusion-prone project.
    - In the company, the SEPG noticed that several projects tend to be in somehow uncontrollable state for a certain period of their development. **(Risky project)**
    - Most of them return to controllable state.
    - But some rarely fall into a really dangerous state. **(Confused project)**

All projects

Risky projects

Confused projects

Evaluation of Risky Projects

- Risky project is evaluated by the SEPG at the end of the development using the quantitative metrics.

  - Using the errors of estimation for cost and duration.

  - If actual cost and duration exceed estimated ones by a certain amount, the project is determined risky.

  - All of the projects in this company can be classified as risky or not.

Outline

- Background and objective

- Definition of "risky" project

- **Our approach**

- Design of a questionnaire (Step 1)

- Construction of a statistical model (Step 3)

- Risk prediction (Step 5)

Characterization of Risky Projects

Phase 1: We design a questionnaire and construct a logistic model to characterize the risky projects.

- **Questionnaire:**
  Extract the intuition from project managers.

- **Logistic model:**
  Analyze the responses objectively.

Phase 2: Using the constructed model, we try to predict the result of new projects.

Outline of Phase 1

Step 1: Design of a questionnaire

Questionnaire

Step 2: Fill in the questionnaire

Responses from managers

Step 3: Construct a logistic model

Phase 2

Project managers

SEPG & Osaka univ. June 8, 2000

ICSE2000 · Limerick · Ireland

Outline of Phase 2

Phase 1

Constructed model at Phase 1

$$E(Y \mid x_1, x_2) = \frac{e^{-5.25 + 2.72 x_1 + 3.98 x_2}}{1 + e^{-5.25 + 2.72 x_1 + 3.98 x_2}}$$

Step 4: Fill in the questionnaire

Responses from managers

Step 5: Calculate the probability

SEPG & Osaka univ.

Project managers of new projects

Outline

- Background and objective

- Definition of "risky" project

- Our approach

- **Design of a questionnaire** (Step 1)

- Construction of a statistical model (Step 3)

- Risk prediction (Step 5)

Design of Questionnaire (Step 1)

- The questionnaire is designed using the risk management manuals in the company.
  - Thus, it includes long-term experience of the software development.

- Characteristics of the questionnaire:
  - Five viewpoints for risk factors are included*.
  - It can be filled in by (at least) the end of the design.

  * Similar views are proposed by SRE model at SEI.

Five Viewpoints for Risk Factors

- ## Requirements
  - Requirement definitions and translation of them into specification.

- ## Estimation
  - Estimations and technical methods for carrying out the estimation.

- ## Team organization
  - Staffing of the project and the fundamental skills of developers.

- ## Planning capability
  - Planning or scheduling method and the resulted project plan.

- ## Project management activity
  - Project management activities during the development.

Questionnaire (Overview)

1.	Requirements	
	1.1	Unreasonable customers.
	1.2	Developers could not elucidate sufficient requirements.
	1.3	Developers misunderstood the requirements of the customer.
	1.4	Lack of interactive agreement regarding requirement specifications between the customer and the developer.
2.	Estimations	
	2.1	There were missing items to be estimated; these items were included in the implicit requirements.
	2.2	The importance of estimations was not well recognized.
	2.3	Non-technical pressure rendered estimates of costs and/or schedules unrealistic.
	2.4	Over-optimism in estimating technical issues.
	2.5	Insufficient estimations were carried out using the results of successful projects in the past.
3.	Team Organization	
	3.1	Wrong people available (lack of skills, lack of training, lack of expertise).
	3.2	Incorrect staffing (too few people for current task).
4.	Planning Capability	
	4.1	Unclear responsibilities and authorities.
	4.2	Inadequate specifications regarding the work product.
	4.3	Inadequate or excessive planning or scheduling of the review process.
	4.4	Lack of commitment on the part of all of the developers with regard to the project plan.
	4.5	Lack of review for the project plan by senior managers.
	4.6	Inadequate control of the development process.
5.	Project Management Activities	
	5.1	Lack of risk management on technical matters.
	5.2	Low morale on the part of the developers.
	5.3	Lack of perception on the part of the managers to ensure a concerned effort.
	5.4	Requirement or specification changes were not managed sufficiently.
	5.5	Lack of progress reporting.
	5.6	Lack of data needed to keep track of a project.

Detail of Questionnaire

2. **Estimation**		
	2.1	There were missing items to be estimated; these items were included in the implicit requirements.
	2.2	The importance of estimations was not well recognized.
	2.3	Non-technical pressure rendered estimates of costs and/or schedules unrealistic.
	2.4	Over-optimism in estimating technical issues.
	2.5	Insufficient estimations were carried out using the results of successful projects in the past.

For each item, one of the four answers "Strongly agree", "Agree", "Disagree" or "Neither agree nor disagree", should be returned.

Outline

- Background and objective

- Definition of "risky" project

- Our approach

- Design of a questionnaire (Step 1)

- **Construction of a statistical model** (Step 3)

- Risk prediction (Step 5)

Construction of the model (Step 3)

- Target projects: 32 projects in 1996 and 1997.

    – These projects were done in similar organizations in the company.

    – We delivered the questionnaire to the project managers of these projects.

    – Based on the responses, we estimate the coefficients in the model.

Evaluation of Projects by SEPG

- Since the projects in 1996 and 1997 finished their development, the SEPG has already evaluated whether they are risky or not.

  – Risky: 10
  – No problem: 22

  - Note that these 32 projects are not all the projects in the company.

Projects in '96 and '97	Evaluation by SEPG
PJ1	Risky
PJ2	Risky
PJ3	Risky
PJ4	Risky
PJ5	Risky
PJ6	Risky
PJ7	Risky
PJ8	Risky
....	....
PJ31	No problem
PJ32	No problem

June 9, 2000

Conversion of Source Data

- For each answer to the item, we assign the following points:

Answer	Point
Strongly Agree	2
Agree	1
Disagree	0
Neither agree nor disagree	1

– We assigned point 1 to "Neither agree nor disagree" because the fact the manager cannot say "disagree" implies some problem in the project.

Conversion of Source Data (cont'd)

- We summed up and normalized these points for each viewpoint (risk factor).

2. Estimation			Evaluation
	2.1	There were missing items to be estimated; these items were included in the implicit requirements.	
	2.2	The importance of estimations was not well recognized.	
	2.3	Non-technical pressure rendered estimates of costs and/or schedules unrealistic.	0.8
	2.4	Over-optimism in estimating technical issues.	
	2.5	Insufficient estimations were carried out using the results of successful projects in the past.	

Source Data from Projects ('96 & '97)

Projects in '96 and '97	Requirements	Estimation	Team organization	Planning capability	Project management activity	Evaluation by SEPG
PJ1	1.50	1.00	0.50	0.83	1.00	Risky
PJ2	1.50	1.00	0.50	1.17	1.00	Risky
PJ3	0.75	1.40	1.00	0.50	0.17	Risky
PJ4	1.50	0.40	1.00	0.83	0.83	Risky
PJ5	1.00	1.00	0.50	0.50	0.83	Risky
PJ6	0.75	0.00	0.00	1.00	0.33	Risky
PJ7	1.50	1.40	2.00	1.67	1.17	Risky
PJ8	1.75	1.80	2.00	1.83	1.33	Risky
....	....	....	....	....	....	....
PJ31	0.00	0.00	0.00	0.17	0.00	No problem
PJ32	0.00	1.20	0.50	0.17	0.33	No problem

Logistic Regression Model

$$E(Y \mid x_1, x_2, \dots, x_n) = \frac{e^{b_0 + b_1 x_1 + b_2 x_2 + \dots + b_n x_n}}{1 + e^{b_0 + b_1 x_1 + b_2 x_2 + \dots + b_n x_n}}$$

- Objective variable Y: Risky(1) or not(0).
- Explanatory variables x_1, ..., x_n: Risk factors.
  - $E(Y \mid x_1, \dots, x_n)$ denotes the conditional probability that a project becomes risky.

- We must estimate stepwisely the values of coefficients $(b_0, \dots, b_n)$ using the source data.

Prediction Model

- Two risk factors "Estimation" and "Planning Capability" are adopted in the model as x_1 and x_2, respectively.

$$E(Y \mid x_1, x_2) = \frac{e^{-5.25 + 2.72 x_1 + 3.98 x_2}}{1 + e^{-5.25 + 2.72 x_1 + 3.98 x_2}}$$

- Result of application to the source projects.

Actual	Predicted	
	No problem	Risky
No problem	18	4
Risky	1	9

27 projects out of 32 were predicted correctly.

(The threshold between risky or not is selected as 0.3)

Outline

- Background and objective

- Definition of "risky" project

- Our approach

- Design of a questionnaire (Step 1)

- Construction of a statistical model (Step 3)

- **Risk prediction** (Step 5)

Prediction of Risky Projects (Step 5)

- By applying the prediction model $E(Y|x_1,x_2)$, we tried to predict the result of 8 projects in 1998.

    – These projects also have almost the same property as the projects used in construction of the model.

    – Since the development of these projects have already finished, we can compare the result of prediction with the result of actual evaluation by the SEPG.

Outline of Phase 2

Phase 1

Constructed model
from 1996 & 1997 at Phase 1

$$E(Y \mid x_1, x_2) = \frac{e^{-5.25 + 2.72x_1 + 3.98x_2}}{1 + e^{-5.25 + 2.72x_1 + 3.98x_2}}$$

Step 4: Fill in the questionnaire

Responses from managers

Step 5: Calculate the probability

SEPG & Osaka univ.

**Project managers
of projects in 1998.**

Result of Experiment ('98)

Projects in '98	Estimation (x_1)	Planning Capability (x_2)	$E(Y|x_1, x_2)$	Actual Evaluation by SEPG
PJ33	0.60	2.00	0.99	Risky
PJ34	1.20	1.33	0.97	Risky
PJ35	1.20	0.33	0.34	Risky
PJ36	0.00	0.17	0.01	No problem
PJ37	0.40	0.33	0.06	No problem
PJ38	0.20	0.33	0.03	No problem
PJ39	0.00	0.17	0.01	No problem
PJ40	0.40	0.67	0.18	No problem

- PJ33, 34, 35 are considered to be risky.
  (Again, the threshold of risky or not is assumed as 0.3.)

- Results of actual evaluation are the same as the results of prediction by the model.

Conclusion

- We proposed a new approach to predict the risky projects based on the questionnaire and statistical model.

  – Experimental evaluation shows that the proposed approach has high predicting capability.

Future Work

  – Apply to the ongoing projects.

  – Improve the design of questionnaire.

  – How to mitigate the risk.