Software project risk probability assessment based on dynamic Bayesian network

Zhang Junguang1, a, Guo Lihong2, b, Xu Zhenchao2, c

1 School of Economics and Management, University of science and Technology Beijing, Beijing, 100083, China

2 School of Economics and Management, University of science and Technology Beijing, Beijing, 100083, China

aemail: zhangjg@ustb.edu.cn, bemail: pmglh10@163.com, cemail: xuzcdsx@163.com

Abstract

Traditional Bayesian network (BN) can only have static analysis which could not reflect the impact of time factors on project risk adequately. For this reason, a software project risk probability assessment model based on dynamic Bayesian network (DBN) is proposed, which combines time series theory and Bayesian theory together to express the risk factor status change relationship between different time segments through probability and directed acyclic graph. Moreover, in the case of lack of sample data, using Leaky Noisy-or gate model to calculate the conditional probability of the nodes will come to a more objective evaluation result. Compared with the assessment results of static Bayesian network (SBN), dynamic Bayesian assessment model improves the accuracy of risk probability assessment of software projects, and provides a more scientific basis for risk control.

Keywords: Project Management; Static Bayesian Network; Dynamic Bayesian Network; Software Project; Risk Probability Assessment

Literature Review

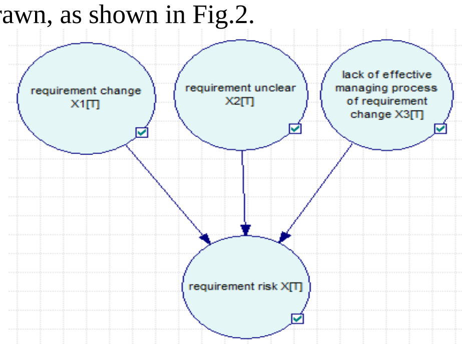
BN represents a useful formalism in the risk analyses domain due to their ability to model probabilistic data with dependencies between events [1].Tang Aiguo et al. combine BN with software project risk assessment to provide a newly evaluating method [2].Jin Junli et al. combine genetic algorithm and EM algorithm with BN to effectively reduce human subjectivity in building BN and its parameters, which makes software project risk assessment more scientific and reasonable [3]. DBN is an extension of BN. DBN introduces time series theory on the basis of SBN and the probability network structure, and then forms a newly stochastic model which is able to deal with time series data.

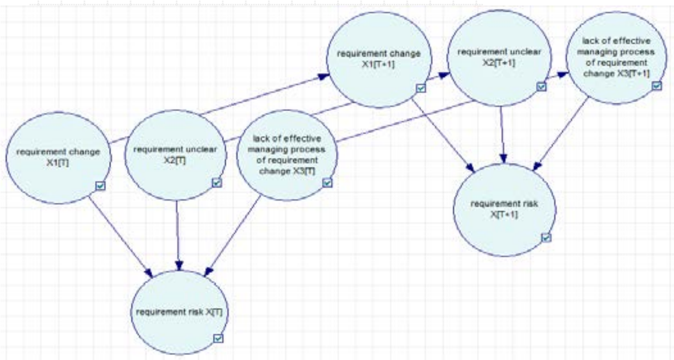
Ann E. Nicholson and M. Julia Flores combine DBN with status and transition models to apply to pasture management. Moreover, they take time factors into account, so that the nature limitations in the domain could be avoided [4].

Eunchang presents a scheme for large engineering project risk management using a Bayesian belief network, but the limitations of his study were the reliance on an expert survey to construct the Bayesian belief network and the consequent requirement for a great effort for data collection [5]. Olga and Daniel used DBN for probabilistic modeling of tunnel excavation processes, which facilitates the quantification of uncertainties in the construction process and of the risk from extraordinary events that cause severe delays and damages [6].

Dynamic Bayesian Risk Probability Assessment Model

The first step in risk assessment is to identify project risks and risk factors. The occurrence probability of requirement risk is high in all the four stages during the project progress, so we will take the requirement risk probability assessment of software projects as an example to demonstrate the feasibility of dynamic Bayesian risk assessment model. We set X to represent requirement risk, the influence factors of which include: requirement change risk X1, requirement unclear risk X2, and lack of effective requirement change control process X3. Firstly, build the priori network Bo of DBN, which is also a SBN. It is shown in Fig.1. After getting Priori network Bo, the corresponding transfer network B– is drawn, as shown in Fig.2.





Fg.1 Priori Network of DBN Fig.2 Transfer Network B– of DB

There are two states at each node: state0=1 represents the state when the risk occurs and state1=0 is on the contrary.

After investigating an software enterprise, we get the data of 100 projects implemented, 70 projects of which have requirement changes during the conceptual stage, then we could get the relevant priori probability is 70/100 = 0.7. By historical statistics, we give a priori probability on each node as follows:

P(X1=1) =0.7, P(X2=1) =0.6, P(X3=1) =0.4.

And the following conditional probabilities are also known:

P(X=1｜X1=1)=0.7, P(X=1｜X2=1)=0.8, P(X=1｜X3=1)=0.6,

P(X=1｜X1=0)=0.5, P(X=1｜X2=0)=0.6, P(X=1｜X3=0)=0.4.

We can obtain the conditional probability parameters of requirement risk for node X, as shown in Table 1. Where Φ is the collection of all the unknown factors except the already know factors, X1 and X2.

Table 1. Conditional probability parameters of requirement risk for node X

Node X	Φ	X_1	X_2
$P(X1=1)$	0.100	0.460	0.550
$P(X1=0)$	0.900	0.540	0.450
Node X	X_3	X_1,X_2	X_1,X_3
$P(X_1=1)$	0.400	0.730	0.640
$P(X_1=0)$	0.600	0.270	0.360
Node X	X_2,X_3	X_1,X_2,X_3	
$P(X_1=1)$	0.700	0.820	
$P(X_1=0)$	0.300	0.180	

We input the obtained priori probability and conditional probability parameters of each node into the model to generate priori network Bo of dynamic Bayesian risk assessment. After running the program, we can get the requirement risk occurrence probability P (X [T]) of the initial state node X, as shown in Fig.3.

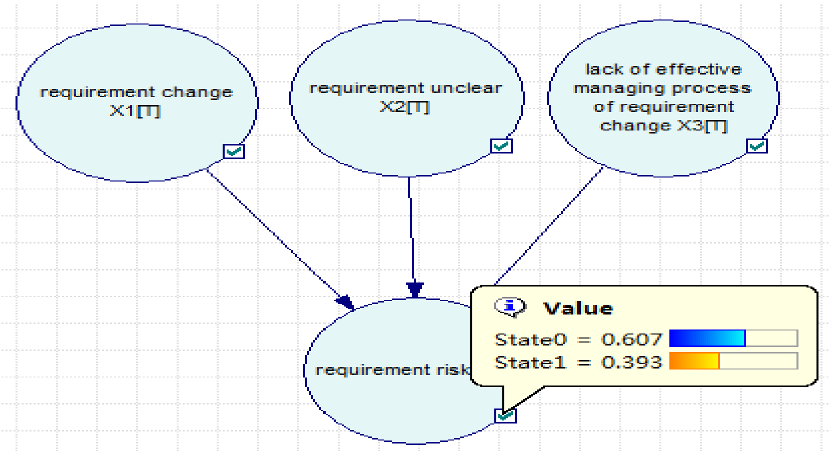


Fig.3 Probability P(X [T]) of Requirement Risk

As the project progresses, there are some changes happen in the designing stage. After re-evaluation based on historical data, the priori probability of each parent node is given as follows:

P(X1=1) =0.5, P(X2=1) =0.4, P(X3=1) = 0.4.

First, the proposed method calculates the transition probabilities between the nodes. Transition probabilities can be obtained: iB–) =(0.2277, 0.7723). The priori probabilities of the three parent nodes in this stage can then be obtained. As shown in Fig.4, Fig.5 and Fig.6.

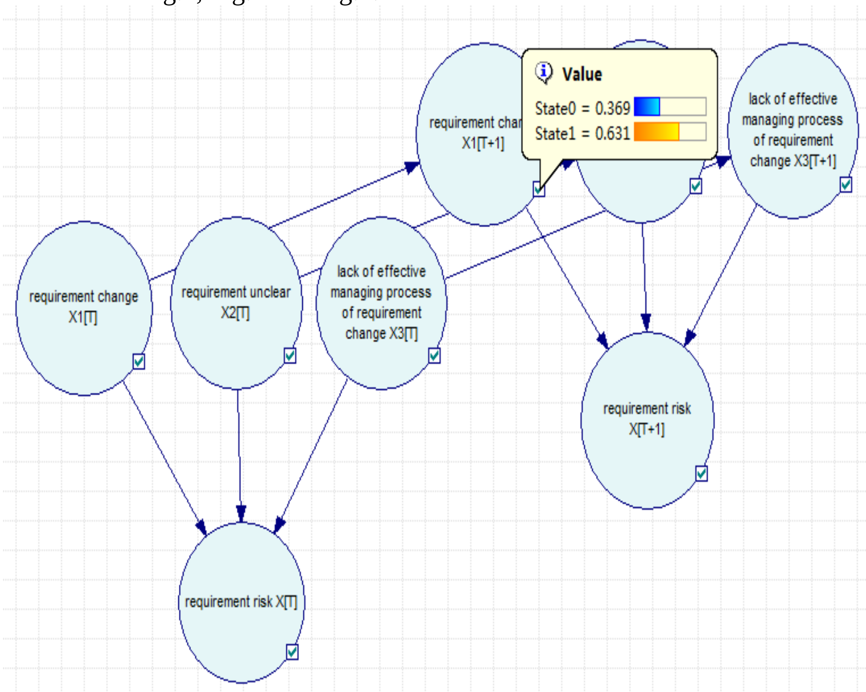


Fig.4 Probability P(X1[T+1]) of Requirement Change Risk

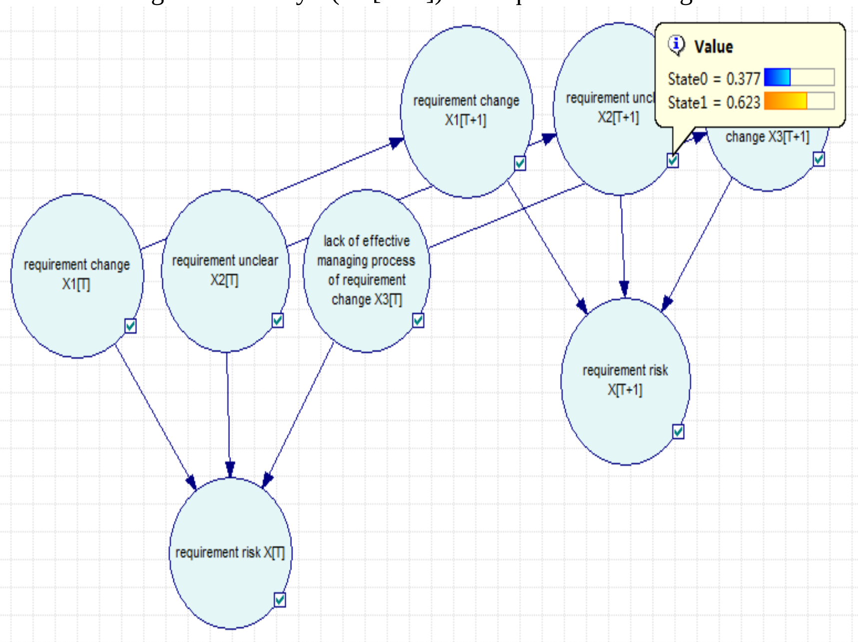


Fig.5 Probability P(X2[T+1]) of Requirements-unclear Risk

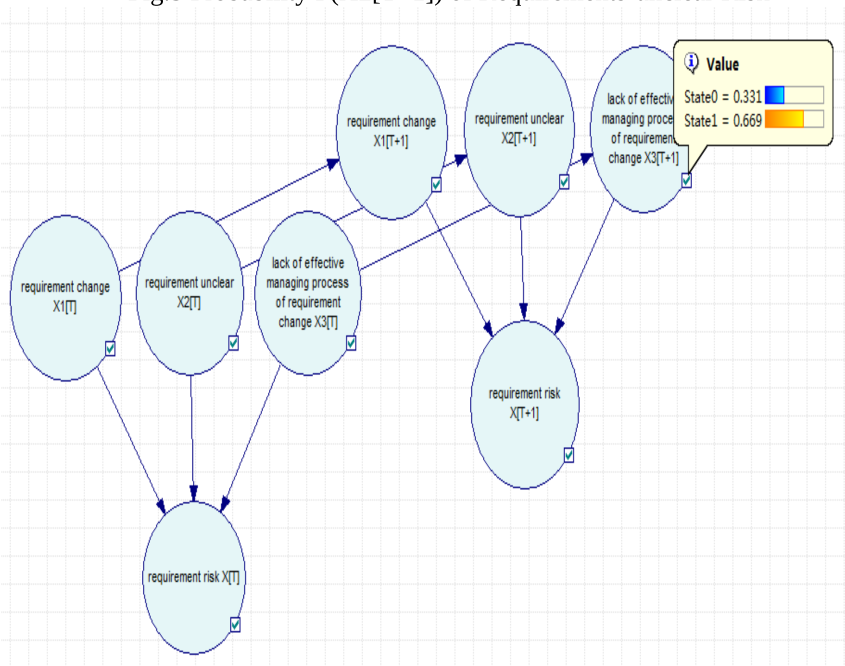


Fig.6 Probability P(X3[T+1]) of Requirement Changing Management Risk

A key assumption made before building DBN is: the conditional probability process of the adjacent time is stationary which means the conditional probability table of each node at every time and the transition probability between different times will not change with time.

The proposed method inputs the priori information and conditional probability parameters of parent node got by deduction into dynamic Bayesian risk assessment transition network B–, and then the occurrence probability of requirement risk during the designing stage will be obtained, as shown in Fig.7.

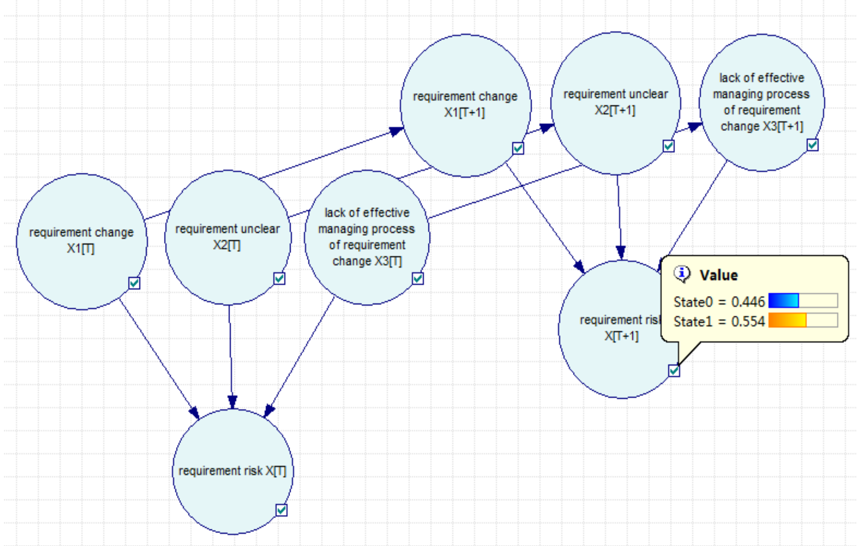


Fig.7 Probability P(X [T+1]) of Requirement Risk

In accordance with the method described above, we can infer the occurrence probability of requirement risk during the four stages, as shown in Table 2.

Table 2 Results of Dynamic Bayesian Probability Risk Assessment

	Conceptual Stage	Designing Stage	Implementing Stage	Finishing Stage
State0	0.607	0.446	0.428	0.356
State1	0.393	0.554	0.572	0.644

Conclusions

This paper puts forward a software project risk assessment model based on DBN. Through the comparative analysis of the results with static assessment model, we find that the dynamic model can effectively analyze project risks, and can get a more accurate assessment result. It improves the risk probability assessment accuracy and provides a more scientific basis for the project manager to control project risks.

As a new project risk probability assessment idea, there has been much progress in DBN risk assessment area. However there are still several limitations that have to be taken into consideration. During the process of research and application, the actual data for risk assessment of software projects can't always be obtained. Even if the data can be obtained, they may be incomplete. Therefore, in future research we will try to get more data from the enterprises by investigation. On the other hand, we intend to improve the algorithm to resolve the problem of incomplete data, and combine Gibbs sampling and Monte Carlo simulation method with dynamic Bayesian model to implement software project risk assessment in case of incomplete data.

References

[1] P. Weber, G. Medina, C. Simon, B. Iung. Overview on Bayesian networks applications for dependability, risk analysis and maintenance areas. Engineering Applications of Artificial Intelligence, 2008, 25(4): 671–682.

[2] Tang Aiguo, Zhao Chuan. Application of Bayesian Networks in Software Project Risk Assessment. Computer Engineering and Application, 2010, 46(7):62-65.

[3] Jin Junli, Zhao Chuan, Yang Jie. Research and Application of Improving Bayesian Network Based on EM-GA. Research on Application of Computer, 2010, 27（4）:1360-1362.

[4] Ann E. Nicholson, M. Julia Flores. Combining state and transition models with dynamic Bayesian networks. Ecological Modeling, 2011, 222(3): 555-566.

[5] Eunchang Lee, Yongtae Park, Jong Gye Shin. Large engineering project risk management using a Bayesian belief network. Expert Systems with Applications, 2009, 36(3): 5880–5887.

[6] Olga Špačková, Daniel Straub. Dynamic Bayesian Network for Probabilistic Modeling of Tunnel Excavation Processes. Computer-Aided Civil and Infrastructure Engineering, 2013, 28(1):1-21.