CHAPTER VI

CONCLUSIONS

6.1 Conclusions

One of the most widely recognized distinction in uncertainty types is between aleatory and epistemic uncertainty. The term aleatory uncertainty is used to describe the inherent variation associated with the risk assessment and project scheduling. Sources of aleatory uncertainty is singled out from other contributors to uncertainty by their representation as randomly distributed quantities that can take on values in an established range of risk variables and temporal variable, but for which the exact value will vary by chance from time to time. The mathematical representation used for aleatory uncertainty is a probability distribution.

This study develops a simulation model to propagate the distributions of risk variables and temporal variable through the simulation model. This is because simulation is widely used to measure the effects of risk factors, such as unforeseeable underground condition, adverse weather, lacking practical experience, failing to distinguish site characteristics, varying from location to location, and being of poor management quality, by introducing randomness into the analysis of construction processes. The developed simulation model uses marginal probability distributions to represent the characteristics of the scheduling variables being modeled. Some assumptions have to be provided for applying probabilistic distributions. For example, the past performance of the same activity has been investigated. Marginal distributions can be established from the investigations which are later used to represent different conditions and random events found in new projects. In practice, however, it is generally known that construction projects are unique. Working conditions for particular projects at the activity level are different. Data related to the risk variables and the temporal variable are usually insufficient and unavailable. Scheduling experts have to use their own judgment based on their experience to assess risk factors and evaluate their impacts on the activity duration. Results certainly involve another type of uncertainties which cannot be addressed by using the probability theory.

The term epistemic uncertainty is used to describe any lack of knowledge or information in any project activity. Incomplete information or incomplete knowledge of some risk factors affecting a considered activity is the fundamental cause of the epistemic uncertainty. Examples of source of epistemic uncertainty associated with the risk assessment and project scheduling are: when there is little or no experimental data for risk variables and temporal variable, limited understanding of complex construction process, and occurrence of risk factors not identified for inclusion in the risk assessment. The mathematical representation of epistemic uncertainty has proven to be much more challenge comparing to the aleatory uncertainty. This study develops a method providing the representation, aggregation, and propagation of the epistemic uncertainty, as well as mixture of epistemic and aleatory uncertainty. The proposed method is designed to reflect the fundamental causes of epistemic and aleatory uncertainty.

Fuzzy set theory is used to address the epistemic uncertainty stemming from systematic and unknown contributions. Probability theory and fuzzy set theory are usually used separately which have been proven inadequate to address every uncertainty (aleatory and epistemic uncertainty). The major limitations associated with the determination of the random, systematic, and unknown contributions can be attributed to the inabilities to graphically present particular uncertainties and to mathematically examine values of every uncertainty.

A new network calculation method accounting for every uncertainty in the risk assessment and project scheduling is developed to undertake the above shortcomings. The proposed method called RAIRFNET aims at providing practical solutions based on the application of appropriate theories to the problems in the risk assessment and the project scheduling. The proposed method incorporates the probability theory and fuzzy set theory to appropriately modeling aleatory and epistemic uncertainty associated with the risk assessment and the project scheduling. The probability distribution presenting variation in activity duration is developed with a consideration about information obtained from the risk assessment. The probability–possibility transformation methods including the Salicone's method and the neurofuzzy metamodel are used to develop the membership function representing a random part of activity duration. Next, the values of uncertainty due to systematic and unknown contributions which present the fuzzy part of activity duration are examined. The internal membership functions (i.e., nil, rectangular, trapezoidal, and triangular) are used to represent the fuzzy part of activity duration. The membership function of activity duration is developed by inserting the internal membership function into the membership function representing the random part. The established membership functions are used in the network calculation performed based on the mathematics for random–fuzzy variables. The calculated project duration is represented by the membership function of which the shape depends on the shape of the membership functions of activity duration.

The performance of the proposed methods (i.e., RAIRFNET using Salicone's method to develop the membership function of activity duration, RAIRFNET using neurofuzzy metamodel to develop the membership function of activity duration, Salicone's method processed on data associated with project completion times, and the neurofuzzy metamodel trained on data associated with project completion times) is tested against simulation by considering the project completion times produced by these methods. The results indicate that the realization of the estimated project completion times can be obtained from the applications of the RAIRFNET as the random–fuzzy durations seem to be able to explain either the aleatory uncertainty associated with the random part or the epistemic uncertainty associated with the fuzzy part of activity duration that cannot be particularly determined by the probabilistic methods. In addition to the use of the random–fuzzy numbers to represent duration, the application of mathematics for random–fuzzy variables to perform the network calculation can appropriately analyze uncertainties due to the random, systematic, and unknown contributions, simultaneously.

The demonstrative applications of the proposed methods to model the uncertainties associated with temporal variables such as the activity duration are provided in this study. It is applicable when the activity duration cannot be particularly specified as a deterministic variable, probabilistic random variable, or fuzzy variable, but it is realistically expressed as a random–fuzzy variable. The integration between the fuzzy set theory and probability theory is provided to model human perceptions together with randomness about real–life problems. For the determination of the random contribution to the aleatory uncertainty associated with activity duration, sufficient data are required by the statistic methods based on probability theory. Normally, historical data are used to develop the probability distribution of duration of project activities. As construction projects are unique, the historical data is inapplicable to a new project. The adjustment of historical data based on human subjectivity is necessary. Imprecision of temporal variable values resulting from the use of human subjectivity is suitably represented by the linear approximation such as the triangular and trapezoidal fuzzy numbers expressed in the quadruples form. The degree of belief of temporal variables typically found in construction scheduling can be represented in a simply and effective manner by a membership function. In this thesis, the temporal variables are considered as the random–fuzzy variables. The integration between the membership function representing uncertainty due to systematic and unknown contributions resulting from human subjectivity and the membership function representing uncertainty due to the random contribution is provided to build the membership function of the random–fuzzy variable. The use of the membership function to represent duration does not only determine every uncertainty involved in the temporal variable, but also accurately facilitate network calculation.

The membership function of the project completion time is alternatively developed by using the Salicone's method and the neurofuzzy metamodel in order to take an advantage of the simulation data associated with the project completion times derived from different types of probability distributions. The credibility coefficient used as a measurement has been proven effective to help the scheduler in interpreting the calculated random–fuzzy results in a meaningful way. As the credibility coefficients can be expressed in the form of the membership function, it can be used in the fuzzy inference, whenever more complex decision making is required. To enable the comparison between the proposed methods and the probabilistic scheduling methods including Monte Carlo simulation, the credibility coefficients are used to compare the performance of these methods.

The results produced by the RAIRFNET considering only the random part are found to be reasonably close to those obtained using Monte Carlo simulation as both of them specifically consider the random contribution to the aleatory uncertainty. The calculation performed using the proposed RAIRFNET requires only one time calculation to provide a solution, while the simulation needs a number of calculations (simulation replications) to provide sufficient data required by the statistical analysis. The proposed RAIRFNET provides the results that consider epistemic and aleatory uncertainty due to the systematic and unknown contribution in the development of the membership function of activity duration, which cannot be determined by simulation. The fact that RAIRFNET is able to determine the aleatory and epistemic uncertainty associated with activity duration and the network calculation, makes it more theoretically accurate than other methods.

In this study, the probability–possibility transformation methods used to develop the membership function representing the aleatory uncertainty due to the random contribution are composed of the Salicone's method and the neurofuzzy metamodel. The spread of activity duration and the project completion times obtained from the Salicone's method depends on means and standard deviations which are obtained from the Monte Carlo simulation. The neurofuzzy metamodel on the other hand provides the spreads of activity duration and project completion times based on the simulation data learnt by a knowledge network. The membership function developed by using the Salicone's method is affected by the central limit theorem, which theoretically compensates the random contribution. The neurofuzzy metamodel constructs the membership function without the compensation of the random effect and the ignorance of the fuzzy effect.

The demonstrative applications of the proposed method are provided. The effectiveness and practicality are presented. It has been seen that the proposed methods can be used as tools for improving and facilitating the creation of schedules taking into account every uncertainty in project environment. These methods provide the reasonable way to the integration of risk assessment into the project scheduling for realistically estimating the project completion time and offer much flexibility in interpreting the calculated random–fuzzy results in a simply manner. The schedule built using this approach can be used in the timely project management.

6.2 Application of Proposed Methods

To illustrate the feasibility of using the proposed methods, the proposed methods are applied to case studies including a bored pile construction of a high-rise building project and a tunneling project carried out in the center of Bangkok. Considering the application, at least two benefits can be achieved: 1) the case study can expose the capabilities of the proposed method and can show the process, step-by-step, for quantifying the duration affected by risk. Additionally, the lessons learnt from the model application can be used for further refinement and improvement of the current methods, and 2) practitioners who work on construction sites or projects can use the proposed method to provide a comprehensive estimate of duration of project activities based on impacts of risk on duration of project activities for any given likelihood of occurrence and consequence of risk factors. The proactive measures will be implemented to prevent risk which might influence the project achievement.

The application of the proposed methods is based on available data. This research acquires data through four approaches. At the beginning of the construction, data related to an existing project are usually unavailable. The historical record and experts' questionnaire become the data acquiring approach that can be adopted to perform the risk assessment and duration estimate at the early stage of construction.

The distributions of temporal variables are initially developed based on historical data adjusted by subjective data. To gather the historical data including a number of units of a product of which the construction durations are extended from the originally estimated duration, and also the extended duration, a collection of historical records kept on the completed projects and held in the head office is required from the construction company. This historical data are entered into the evaluation equations. Subjective data are required for adjusting the quantitative data acquired through the historical records. The expert's questionnaire is used to collect subjective data. The subjective risk assessment is performed based on the available information associated with an existing project. The experts are required to answer the subjective questions based on given definitions of attributes of a risk factor. When the construction is proceeding, the factual data become available. The availability and also quality of factual data depends on the progress of the construction.

At the early stage of construction, it is not possible to attain data associated with each production unit as the construction is not yet carried out. During construction, such data might be unattainable because of the dynamic and complex nature of a construction project. To gather the required data, experts were asked to identify risk factors having significant impacts on each activity based on information probably obtained from the early stage of the construction at which the direct observations can be made. The subjective risk assessment is performed by using the direct observations in conjunction with the designed experts' questionnaire. The experts were asked to complete the questionnaires. The experts have to answer the subjective questions in the questionnaire based on given definitions of attributes of a risk factor. The results are the likelihood of occurrence and consequence of a particular risk factor.

As the factual data are insufficient to perform the proposed methods at the early stage of the construction and probably construction stage, the historical record and experts' questionnaire are used to provide data required to perform the risk assessment and duration estimate. The output is duration of a particular activity estimated by accounting for currently assessed values of attributes of the identified risk factors. The probability distributions representing uncertainty due to a random contribution involved in the estimated duration of project activities and the assessed values of attributes of a particular risk factor are then developed. The established distributions are fed to the simulation process so as to estimate a project completion time based on the simulated durations of project activities. Thus, the project completion time determined only uncertainty due to a random contribution is obtained by executing the simulation.

There is another type of uncertainty involved in durations of project activities estimated by considering impacts of risk factors, uncertainty. This type of uncertainty is brought about by total ignorance and unknown, or uncompensated, systematic contributions involved in the assessed values of attributes of risk factors and the estimated duration. A membership function representing the systematic and unknown contributions should be built. In this way, the distribution of the possible durations is described in terms of a probability distribution and framed within the probability theory, if the contributions to uncertainty are random contributions. The distribution of the possible durations is described in terms of a possibility distribution or a membership function and framed within the fuzzy theory, if the contributions to uncertainty are not random. However, in practice, all different kinds of contributions to uncertainty are simultaneously affecting the duration of project activities and must be then considered together. The membership function representing a nonrandom contribution should be considered together with a membership function representing a random contribution transformed from the corresponding probability distributions. Although perfectly coherent from a strict theoretical point of view would be impractical, the aim of this research is to apply a single mathematical object able to consider these contributions. The durations of project activities are represented by the random–fuzzy variables.

As the durations of project activities are considered as random–fuzzy variables represented by the membership functions, the durations of project activities must be mathematically combined in order to compute the project completion time. Suitable mathematics is used in the network calculation to consider a different behavior for the random and nonrandom parts of durations of project activities. The internal membership function representing the uncertainty due to systematic and unknown contributions must adhere to the mathematics of the intervals. The behavior of the random parts of the random–fuzzy variable must be addressed similar to the behavior of the probability distribution functions when they combine together.

As the durations of project activities are represented by the membership functions, a number of simulation replications are not required to estimate the project completion time. The durations of project activities represented by the membership functions are inputted into the network calculation. The project completion time is computed based on mathematics for random–fuzzy numbers. Only one time calculation is required to estimate the project completion time.

Even though the proposed methods have been developed to consider every uncertainty involved in the risk assessment and duration estimate, there are still some cases that cannot be addressed by the proposed methods. Those are the use of a totally generic random–fuzzy variable to represent the temporal variables and risk variables. Only particular subclass can be used to represent these variables. For example, in the case study different kinds of uncertainty contributions (random and nonrandom) do contribute to the width of all confidence intervals associated with the durations of project activities for every possible level of confidence. It means that the random–fuzzy variables employed to represent the durations of project activities must adhere to the condition that both membership functions are normal. Moreover, from a strict mathematical point of view, a membership function must be convex. Thus, the membership function transformed from the probability distribution function has always a single peak value. It means that the interval of confidence of the random–fuzzy variable corresponding to level $\alpha = 1$ always degenerate into an interval of confidence of the fuzzy variable.

In this research, the rectangular, triangular and trapezoidal internal membership functions which are normal are used to represent uncertainty due to the systematic and unknown contributions, while the triangular membership function which has a single peak value is used to represent uncertainty due to the random contribution. Nevertheless, uncertainty contributions (random and nonrandom) involved in the risk assessment and duration estimate from several real–world projects might be appropriately represented by the membership functions which are not normal. As stated, the bimodal probability distribution functions cannot be used to represent uncertainty due to the random contribution as they cannot be transformed into a membership function. However, in practice, bimodal probability distribution functions can be found in the duration estimate, when the duration is supposed to concentrate toward the edges of the interval within which the durations are supposed to lie. This case can be seen as if two different durations are possible. In this case, the mathematics for random–fuzzy variables cannot be applied. Thus, this case should be considered carefully while performing the risk assessment and duration estimates. Two different durations supposed to lie at the edges of the interval should be treated as two different random–fuzzy variables in order to keep a convex membership function.

6.3 Contributions

The contributions of this study are:

1) The development of the risk assessment model for estimating duration of project activities affected by risk factors. The development of the risk assessment model has considerably improved the accuracy of the estimated duration of project activities.
2) The development of a random–fuzzy network scheduling method for analyzing random and nonrandom contributions, simultaneously. The random – fuzzy network scheduling method undertakes the limitations found in many researches.
3) The development of the risk assessment integrated within the random–fuzzy network scheduling method (RAIRFNET). RAIRFNET integrates the developments of the risk assessment model and the random–fuzzy network scheduling method. It reasonably applies suitable theories to address uncertainties involved in the risk assessment and the network calculation. It also provides a practical means for schedulers to estimate duration of project activities and a project completion time where the duration of project activities is imprecisely known.
4) The development of a neurofuzzy metamodel for estimating duration of project activities affected by risk factors and a neurofuzzy metamodel for estimating project completion time taking into account impacts of risk factors. The relationship between attributes of risk factors and duration is suitably represented by using the nonlinear approximation. In addition, the

neurofuzzy metamodel is computationally inexpensive which can provide a modeling output based on the values of inputs fed into the knowledge network.

By applying the proposed method, the uncertain durations of project activities affected by risk factors and the changed critical paths associated with each unit of a product and the whole project can be examined. With this information, the difference between the analysis results and actual durations could be detected. The difference is caused by several sources of errors including ignorance of particular risk factors and attention on unnecessary risk factors, over-estimation and under-estimation of durations of project activities, and subjective assessment based on human's experience, knowledge, judgement, and biases.

To gain an additional advantage of applying the proposed methods, the obtained results should be used to make a decision about the proper reactive and proactive response measures that should be imposed on the site in order to prevent impacts of risk factors on construction activities. This probably revises the working conditions and construction strategies for the remaining production units. The proposed methods are flexible and adaptive. Additional and updated data can be easily included in the knowledge structure development. Furthermore, the updated predictive models can be employed to assist in the project control purpose as the construction planning and scheduling might be adjusted according to the analysis results.

Based on this application, the proposed methods and the developed predictive models are considered useful and applicable for evaluating impacts of risk factors on duration of project activities under dynamically varying construction conditions and constraints. It also assisted in quantifying impacts of risk factors at the early stage of construction and during construction. The experience obtained from former units of products already made can be employed to update the main database of the project. Therefore, the more reliable predictive model for succeeding production units can be developed. In addition, the lesson learnt from the current project could be included into the knowledge base that would be beneficial for the development of a model for duration estimate for future projects.

6.4 Recommendations for Future Work

The successful demonstrative application of RAIRFNET has been proven adequate to provide realistic schedules by appropriately using the probability theory and the fuzzy set theory to generate network scheduling. However, there are a number of possible developments that could improve performance of the proposed approach which include:

1) The process of risk assessment proposed in this study can be improved by providing more reasonable integration between quantitative and subjective data associated with the attributes of risk factors with respect to available data.
2) The RAIRFNET has successfully applied to particular construction, the bored pile construction of a high-rise building project and the tunneling project. A case study illustration is cited in Chapter IV to demonstrate the ability of the proposed RAIRFNET method for achieving a realistic and interpretable schedule. To better investigate the advantages and disadvantages of the RAIRFNET method, a wider set of applications with high-dimensional data are not yet taken into consideration. The research contribution in terms of generalization can be evaluated by applying the RIARFNET to other types of construction, and other industries which involve large sets of high-dimensional data which are affected by missing, vague, imprecise data and a large number of risk factors.
3) The RAIRFNET method can be used to examine the cost of project activities. The dependencies between random–fuzzy durations and their associated random–fuzzy costs would have to be determined. In addition, the relationship between the cost variables and their risk factors would have to be captured by the knowledge network in order to estimate cost of project activities. The more realistic schedule is generated by performing the random–fuzzy time–cost tradeoff using the mathematic for random–fuzzy variables.