Automation and Robotics in Construction XII
E.Budny, A.McCrea, K.Szymanski (Editors)

FUZZY vs PROBABILISTIC SCHEDULING

Moselhi, O. and Lorterapong, P.

Centre for Building Studies, Concordia University, 1455 de Maisonneuve Blvd. West, Montreal, Quebec, H3G 1M8, Canada.

Abstract

This paper provides a comprehensive comparison between two groups of nondeterministic scheduling methods: 1) probabilistic and 2) fuzzy set based methods. A numerical example is used to demonstrate the differences between the two groups of methods. The comparison is conducted considering three aspects: 1) theoretical assumptions, 2) data acquisition and computational effort, and 3) scheduling information. The results indicate that the fuzzy network method can overcome some of the limitations associated with PERT and Monte Carlo simulation. The fuzzy network method is capable of providing realistic and useful information to the project team.

1. INTRODUCTION

Analytical tools such as the Critical Path Method (CPM) and Precedence Diagram Method (PDM) are being used extensively to analyze construction project networks. Both CPM and PDM methods, however, assume scheduling problem to be deterministic. In real-life operations, however, construction projects are normally executed under uncertain environments. With these uncertainties surrounding activities and resources data, it is unlikely that such deterministic methods can be used effectively. To circumvent such a limitation, two lines of research have been developed. (i.e. proabailistic method [1,5], and fuzzy set based methods [3,6,10,11].

The main objective of this paper is to demonstrate the advantages and limitations associated with these methods, and to compare some of the essential characteristics of the scheduling methods based on fuzzy set theory and those inherent in the widely used techniques of PERT and Monte Carlo simulation. In the following section a brief review of the fundamentals of fuzzy set theory and its applications in network scheduling is presented. Detailed descriptions of fuzzy set theory can be found in the references listed at the end of this paper [4,6,9,13,14].

2. FUZZY SET BASED SCHEDULING METHODS

Fuzzy set theory was developed in the mid 70s by Zadeh in an effort to provide a basis to handle uncertainty that is nonstatistical in nature [13]. Basically, a fuzzy set is a class of objects (u) associated with their respective degrees of membership [μ(u)] within the set. The theory differs from the conventional crisp sets mainly in the degrees by which an object belongs to a set. In the crisp set theory, objects are either included or excluded from a set. In the fuzzy sets theory, on the other hand, objects are described in such a way to allow a gradual transition from being a member of a set to a nonmember. A widely used concept derived from the fuzzy set theory is fuzzy numbers [4,6,9]. A fuzzy number is a continuous fuzzy set that contains two properties: 1) convexity and 2) normality. They are used, for example, to represent imprecise numerical quantities such as "approximately 10 days," "about 8 weeks," etc. Though fuzzy numbers can take various shapes, linear approximations such as triangular and trapezoidal fuzzy numbers are used frequently. A trapezoidal fuzzy number can be represented by the quadruples (a,b,c,d) where a and d are the lower and upper bounds, while b and c are the lower and upper modal values, respectively. The representation using the quadruple form is effective (facilitating computations and providing flexibility). Various degrees of temporal imprecision can be expressed in the quadruples form [11].

In general, arithmetic operations (e.g. addition, subtraction, multiplication, etc.) and set theoretic operations (e.g. union, intersection, etc.) are performed using a technique called the max-min convolution [9]. Previous work in applying fuzzy sets theory to network scheduling utilizes the this technique for calculating the project duration and scheduling events [3]. The major shortcoming of the Max-Min convolution is that it requires excessive computational efforts. Such a limitation has been circumvented by the use of "fuzzy interval", expressed in the quadruples (a,b,c,d) to represent activity durations. Having used the quadruples, some of the laborious arithmetic operations that are based on the Max-Min convolution can be replaced by the following direct operations [4,6].

Let $M = (a_1,b_1,c_1,d_1)$, $N = (a_2,b_2,c_2,d_2)$ be two fuzzy numbers. The direct operations on these numbers are:

$$M \oplus N = (a_1 + a_2, b_1 + b_2, c_1 + c_2, d_1 + d_2) \tag{1}$$

$$\tilde{\max}(M,N) = (\vee(a_1,a_2), \vee(b_1,b_2), \vee(c_1,c_2), \vee(d_1,d_2)) \tag{2}$$

where $\oplus$ represents fuzzy addition, $\vee$ denotes maximum, $\wedge$ represents minimum, $\tilde{\max}$ symbolizes fuzzy maximum, respectively.

Due to some limitations associated with traditional fuzzy arithmetic operations, previous fuzzy network methods do not support the backward pass calculations [3,11] as traditionally being performed in CPM. This major shortcoming has been overcome by the use of fuzzy bounds along with new treatments [11]. The method is called FNET (Fuzzy NETwork Scheduling). FNET has currently been extended to incorporate resource constraints into scheduling [10].

In addition to the results obtained from fuzzy calculations, fuzzy set theory and its derivatives (i.e. the possibility measure, and agreement index concepts) can also be used to assess the possibility of occurrence of expected project events (e.g. expected project completion time). Possibility measure is a concept introduced by Zadeh [14] to evaluate the degree of possibility of a variable pertaining to a certain event. That underlying event could be fuzzy (e.g. a switch-gear will arrive on site about mid of March 1995) or nonfuzzy (i.e. the project completion date is 24 July 1995). Let U be a universe of discourse, X be a fuzzy project duration resulting from fuzzy network analysis, F be a fuzzy expected event, and N be an expected crisp time interval, the compatibility between the two events X and F is calculated as follows:

$$\mathrm{Poss}\{X \text{ is } F\} = \bigvee_{u \in U} [\mu_F(u) \wedge \mu_X(u)] \qquad (3)$$

In the case where the event being examined is a crisp event (i.e. event N), Eq. 3 can be rewritten as:

$$\mathrm{Poss}\{X \in N\} = \bigvee_{u \in N} \mu_X(u) \qquad (4)$$

The agreement index is another useful concept which is analogous to the cumulative probability value [9]. This index measures the ratio of the intersection area between two fuzzy events (i.e. the calculated and expected events) with respect to the area of the calculated event. More specifically, let A and H be any two events being considered. The agreement index of A with respect to H, ι(A,H) is defined as follows:

$$\iota(A,H) = [\mathrm{Area}\ (A \cap H)] / \mathrm{Area}\ A \qquad (5)$$

3. COMPARISON BETWEEN FUZZY AND PROBABILISTIC METHODS

This section presents a comparison between FNET and probabilistic network scheduling methods including PERT and Monte Carlo simulation. The comparison is performed through the use of a numerical example extracted from the literature [9]. A number of scenarios are generated using a network example to illustrate the attractive features of FNET over currently used probabilistic methods. The comparison is made primarily on: 1) theoretical assumptions, 2) data acquisition and computational effort, and 3) scheduling information provided to the user, upon completion of the analysis.

3.1 Theoretical Assumptions

In general, PERT is based on three basic assumptions [2,7,8]: 1) activities are independent, 2) the critical path is substantially longer than other paths, and 3) the critical path contains a "sufficiently large" number of activities. The first assumption also applies for the FNET method, but it is not necessary for Monte Carlo Simulation only when the analysis accounts for the correlations among the project activities. The data required for this type of analysis is difficult to obtain and maintain, and is rarely available in practice.

The second assumption has consistently been criticized as a major drawback for PERT. In the situation where project networks contain several near-critical paths, project durations calculated using PERT are underestimated [2,8]. This is due to the fact that the determination of the critical path considers only the mean of activity durations involved. No consideration, however, is given to the variances. This process can result in a loss of information which occurs particularly at the joint nodes where two or more activities meet. Such a drawback has been treated in FNET. In the forward pass calculation, the expected early start time of a joint node is determined from the largest early finish of all the activities leading to that node. Unlike PERT, the M$\tilde{a}$x operation (Eq.2) employed in FNET performs paired-wise comparisons for each and every element of all the quadruples, and accordingly selects the maximum value of each pair to represent their respective elements for the new quadruples. This procedure thus eliminates that problem.

The third assumption enables the use of the central limit theorem [7]. The central limit theorem imposes a specific constraint on the calculated project duration. It suggests that the project duration resulting from the use of this theorem is normally distributed, disregarding the distribution assumed for activity durations [2]. The theorem, however, requires the marginal distributions associated with each independent variable (i.e. the duration of activities) be identical [12]. This restricts the generality of the method and may limit its applications to real-life problems. FNET, on the contrary, does not impose such a constraint. As such, FNET provides more flexibility and practicality in this regard.

3.2 Data Acquisition and Computational Effort

Probability theory was originally developed in an effort to provide a basis for dealing with uncertainty due to randomness [14]. The theory is conceptually based on experiments which are repeated, theoretically under identical circumstances, and without mutual dependence. To model the characteristic values of a variable (i.e. activity durations in the case of scheduling), historical data must be collected. A frequency histogram of the gathered data is then developed. Next, a probability density function is then subjectively selected to represent the data set. The statistical parameters associated with the selected probability function are estimated. The test for goodness of fit is then conducted to measure the quality of the fit obtained. Theoretically, the above procedure is required for all activities in the network. If the simulation is to be used, correlation coefficients among project activities are additionally required. This process requires considerable documented field observations which are rarely available in practice.

In real-life projects, activity durations can be estimated as being within a certain interval without even any knowledge of a probability distribution within that interval. For example, an expert may say that the duration of excavating for a foundation is generally between 7 to 10 days. But due to other factors which can not be controlled or predicted, the progress may occasionally be as slow as 18 days or as fast as 5 days. This typical situation is more naturally represented by a trapezoidal fuzzy number rather than by a probability distribution. The main advantage of such a representation is that there is no need to force the manager to give a single number nor a probability distribution to represent the activity duration. With respect to the computational requirements, fuzzy arithmetic operations are simple and can easily be performed. The calculations may not be as simple as those used in PERT, but they are more transparent. Manual calculations are possible for small-size networks. FNET provides a direct solution, it requires less computation time than Monte Carlo simulation.

3.3 Scheduling Information

A nine-activity project network extracted from the literature was analyzed [8]. The network and its detailed description are shown in Fig. 1. All three methods (PERT, Monte Carlo simulation, and FNET) are applied. The estimated optimistic, most likely and pessimistic durations for the nine activities are shown in Table 1. Two different probability distributions: the triangular and beta are assumed. In the case of the beta distribution, two different limits [8]: 1) the ninety-fifth percentile values, and 2) the absolute for the optimistic and pessimistic activity durations are assumed. The means and variances of activity durations are also listed in Table 1. To apply FNET, a triangular membership function is assumed in consistency with the data used in PERT and Monte Carlo simulation. Activity durations are represented by fuzzy numbers expressed in the quadruple format (see also Table 1). Equations 1 to 5 are used in FNET.

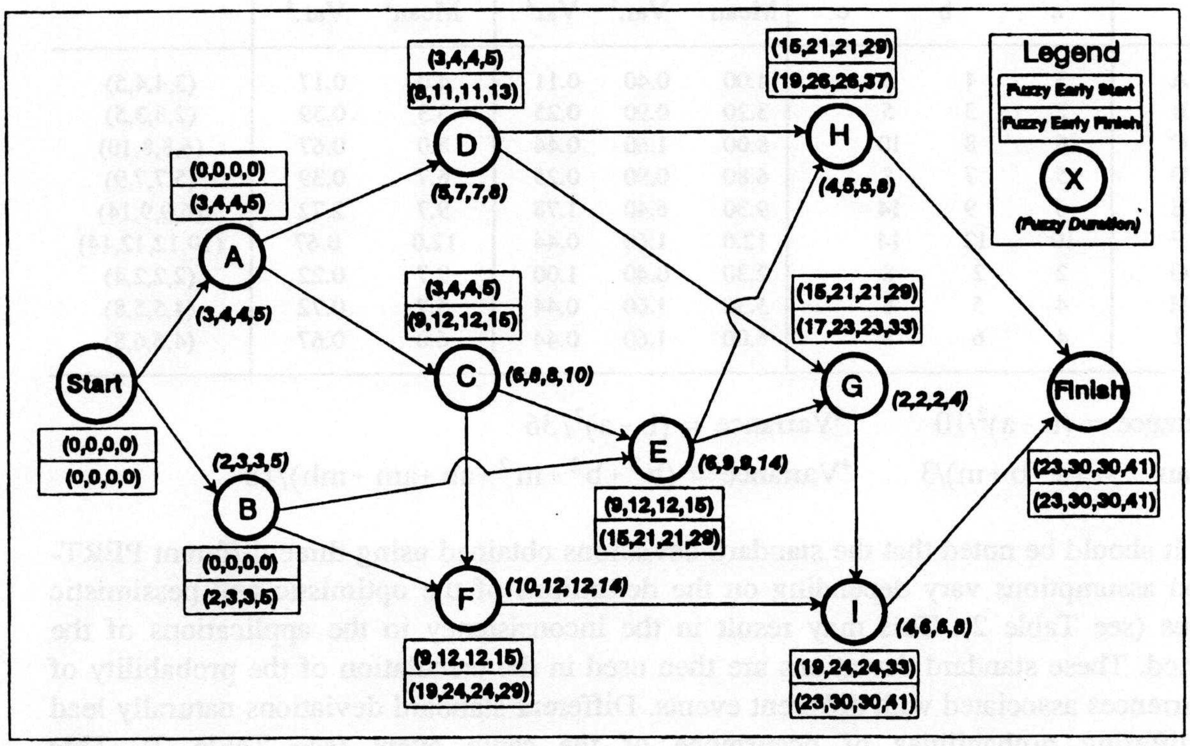

Figure 1. Network Example (adopted from [8])

Table 2 shows the outputs generated by the three methods. Probabilistic methods produce scheduling outputs in terms of means and variances of a probability density function. The distribution of project durations is assumed to be normal in all cases. FNET, on the other hand, produces an output in the form of simple linear approximations. For this example, the fuzzy project completion time is (23,30,30,41). This can be interpreted as *about 30 days with an absolute minimum of 23 and an absolute maximum of 41 days respectively*. The shape of fuzzy project duration produced by FNET depends on those of the input data. For example, if all input of activity durations are triangular, the resulting project duration will be the triangular. This form of output is believed to be easier for the project team to adopt for practical purposes. For example, it is more direct and more natural to represent linguistic interpretations such as "about", "between", etc. in the form of the trapezoidal or triangular fuzzy numbers rather than in the form of a probabilistic density function.

Table 1
Network Input Data

Activity	Estimated Activity Durations			Beta Distribution			Triangular Distribution		Triangular Fuzzy No.
	a	b	c	Mean	Var.[1]	Var[2]	Mean[3]	Var.[4]	
A	3	4	5	4.00	0.40	0.11	4.0	0.17	(3,4,4,5)
B	2	3	5	3.20	0.90	0.25	3.3	0.39	(2,3,3,5)
C	6	8	10	8.00	1.60	0.44	8.0	0.67	(6,8,8,10)
D	5	7	8	6.80	0.90	0.25	6.7	0.39	(5,7,7,9)
E	6	9	14	9.30	6.40	1.78	9.7	2.72	(6,9,9,14)
F	10	12	14	12.0	1.60	0.44	12.0	0.67	(10,12,12,14)
G	2	2	8	2.30	0.40	1.00	2.7	0.22	(2,2,2,4)
H	4	5	8	5.30	1.60	0.44	5.7	0.72	(4,5,5,8)
I	4	6	8	6.00	1.60	0.44	6.0	0.67	(4,6,6,8)

[1]Variance = $(b-a)^2/10$ [2]Variance = $(b-a)^2/36$
[3]Mean = $(a+b+m)/3$ [4]Variance = $(b^2+b^2+m^2+ab+am+mb)/18$

It should be noted that the standard deviations obtained using three different PERT-based assumptions vary depending on the definitions of the optimistic and pessimistic values (see Table 2). This may result in the inconsistency in the applications of the method. These standard deviations are then used in the calculation of the probability of occurrences associated with different events. Different standard deviations naturally lead to differing probabilities of occurrence of the same event (see Table 2). This inconsistency, although not shown in this example, can also occur in Monte Carlo simulation. This shortcoming, however, can be overcome in the FNET method.

The results shown in Table 2 reveal a close correlation between the probability of occurrence of the expected event generated by Monte Carlo simulation with correlations and that obtained using FNET (see the sixth column in Table 2). In general, the possibility of occurrence produced by the fuzzy network method is the most pessimistic. This can be related to the manner in which the uncertainty inherent in the input data is modeled. In FNET, all of the three scenarios (i.e. the optimistic, the pessimistic, and the most possible) associated with the durations of the critical activities are actually included in the spread of the calculated fuzzy project durations. The standard deviation calculated for a duration is usually smaller than the difference between the optimistic and pessimistic values. As such, the spread of fuzzy project durations tend to be larger than the standard deviation obtained from PERT and Monte Carlo simulation. Therefore, the possibility of occurrence calculated using FNET (Eq.5) tends to be more pessimistic than those obtained from PERT and Monte Carlo simulation. The pessimistic value can in fact have positive influences, drawing the attention of the project team to scheduling slippage.

Table 2
Results obtained using different methods

Methods	Input Distribution	Output Distribution	Project Duration (D)	Std.	Prob (D≤35)	Prob (D=35)	P(D=35)> P(D=28)?
PERT[a]	Beta	Normal	30	2.30	0.985	NA	NA
PERT	Triangular	Normal	30	1.47	0.999	NA	NA
PERT[b]	Beta	Normal	30	1.19	1.000	NA	NA
MC[c]	Beta	Normal	30.9	2.50	0.945	NA	NA
MC[d]	Beta	Normal	36.5	4.90	0.799	NA	NA
FNET	Triangular	As input	(23,30,30,41)	NA	0.818	0.55	Yes[e]

[a] The 95th percentile activity durations values
[b] The absolute limit activity duration values
[c] Monte Carlo simulation (Independent)
[d] Monte Carlo Simulation (Correlations) each activities' durations
[e] Poss (D=35) = 0.55, Poss (D=28) = 0.71

In addition to the traditional scheduling information described earlier, the FNET method also provide other useful information which, by definition, can not be obtained from probabilistic methods. The possibility of such an event as the project duration exactly equal to 35 days can easily be calculated using the concept of possibility measure incorporated in FNET. The possibility of having the project duration exactly equal to 35 days was calculated using Eq. 3 to be 0.55 (see Table 4.2). This possibility, however, does not have a direct analogy in the PERT or Monte Carlo Simulation. This is due to the fact that the probability of a strict equality [i.e. P(T = 35)] is always zero when a continuous probability distribution is used. Accordingly, probabilistic methods can not provide an answer to such a question as "*which of the following project durations: exactly 35 days and exactly 28 days is more likely to occur?*". The results obtained using the FNET method reveal that the project duration is more plausibly to be 28 days than 35 days.

4 SUMMARY CONCLUDING REMARKS

For those situations where project teams can not specify activity durations as exact or deterministic numbers, the durations can naturally and realistically be represented as fuzzy numbers.

If the input data are fuzzy then the scheduling outputs should also be fuzzy. This paper demonstrates the use of fuzzy sets theory for modeling uncertainty associated with activity durations in network scheduling. A network example adopted from the literature is worked out to compare the capabilities of the fuzzy network method with the well known probabilistic methods such as PERT and Monte Carlo simulation.

The comparison is carried out on three aspects: 1) theoretical assumptions, 2) data acquisition and computational effort, and 3) scheduling information. With respect to the theoretical assumptions FNET can alleviate a major shortcoming associated with PERT (i.e. focus on a single critical path) through the use of fuzzy maximum operation. Unlike probabilistic methods, FNET does not require the identical probability distribution for all activity durations. As for the data acquisition FNET is more flexible and does not require historical data which, in practice, is difficult to obtain. With respect to the computational requirements, fuzzy arithmetic operations are simple and can easily be performed. The calculations may not be as simple as those used in PERT, but they are more transparent. Manual calculations are possible for small-size networks. Since FNET provides a direct solution, it requires less computational time than Monte Carlo simulation. The FNET method provides natural and more meaningful results than those obtained by PERT and Monte Carlo simulation.

REFERENCES

1. S.M. AbouRizk and D.W. Halpin, "Probabilistic Simulation Studies for Repetitive Construction Processes", J. Constr. Engrg. and Mgmt, ASCE, 116(4), 1990.
2. A.B. Badiru, Quantitative Models for Project Planning, Scheduling and Control, West Port, CT, Quorum Books, 1993.
3. S. Chanas and J. Kamburowski, "The use of fuzzy variables in PERT", Fuzzy Sets and Systems, 5, 1981.
4. S.J Chen and C.L. Hwang, Fuzzy multiple attribute decision making: methods and applications. Springer-Verlag., Berlin, Germany 1992.
5. C.F. Diaz and C. Hadipriono, "Nondeterministic network methods", J. Constr. Engrg. and Mgmt., ASCE, 119(1), 1993.
6. D. Dubois and H. Prade, Possibility theory: An approach to computerized processing of uncertainty. Plenum Press, New York, N.Y., 1988.
7. S.E. Elmaghraby, Activity Networks: Poject Planning and Control, John Wiley & Sons, New York, 1977.
8. C. Hendrickson and T. Au, Project Management for Construction," Prentice Hall, Englewood Cliffs, N.J., 1989.
9. A. Kaufmann and M.M. Gupta, Introduction to fuzzy arithmetic theory and applications. Van Nostrand Reinhold, New York, N.Y., 1985.
10. P. Lorterapong, "A Fuzzy Heuristic Method for Resource-Constrained Project Scheduling", Project Management Journal, December, 1994.
11. P. Lorterapong and O. Moselhi, "Project-Network Scheduling Using Fuzzy Set Theory", Manuscript submitted for possible publication in the ASCE J. Constr. Engrg. and Mgmt., 1995.
12. J.S. Milton and J.C. Arnold, Probability and Statistics in the Engineering and Computer Sciences, McGraw-Hill, New York, N.Y., 1986.
13. L.A. Zadeh, "Fuzzy sets.", Information Control, 8, pp. 338-353, (1965).
14. L.A. Zadeh, "Fuzzy sets as a basis for a theory of possibility", Fuzzy sets and Systems, 1(1), 1978.