Advanced Engineering Technology and Application
An International Journal

http://dx.doi.org/10.18576/aeta/050202

Using Fuzzy Logic to Obtain PERT Three-Time Estimates in Oil and Gas Projects

*Yerkezhan Shakenova**

Kazakh - British Technical University, Almaty, Kazakhstan

Received: 2 Feb. 2016, Revised: 7 Apr. 2016, Accepted: 8 Apr. 2016
Published online: 1 May 2016

Abstract: A lot of projects especially in Oil and Gas sector are out of schedule. Main problem is wrongly calculated activity durations, because Oil and Gas projects are very large and have many activities. In most cases experienced person, who may not be expert in statistics or in its area at all, chooses durations of project activities. Moreover experience and knowledge of person is always limited and his/her opinion is subjective. We propose method of survey for experts in a field of project's area, especially Oil and Gas sector. During the survey experts will be asked to answer questions about activity durations in intervals, not in exact numbers. Survey uses linguistic variables as answers and calculates fuzzy numbers from them. Further defuzzified activity durations will serve as input for simulation.

Keywords: fuzzy logic and sets, project management, cpm, pert three-time estimate, expert

1 Introduction

A project is a temporary endeavor undertaken to create a unique product, service, or result (Pmbok Guide). It contains set of operations designed to realize given goals and obtain desired results. Project can be considered as successfully completed, when all the objectives are achieved within a planned time frame using estimated amount of budget. Projects vary in size and industry. Also one can divide them by time it will take to complete, so there are long-term or short-term projects [1].

In oil and gas industry there are lots of types of projects. Starting with exploration and production, well control, shipping and transportation, storage, refining ending up with marketing and retail sales. These projects nowadays differed from other industries by their complexity. Firstly, in terms of size projects become significantly larger over the past decade or so. Secondly, they are technologically arduous. Because demand for oil products constantly increasing, but oil resources declined in most areas. Therefore oil acquiring is done in more difficult areas. There are several unconventional areas for oil and gas exploration: shale gas, coal seam gas, oil sands, ultra-deepwater, the Arctic, etc [3]. In order to run these kinds of projects companies involve lots of stakeholders and allocate large amount of budget

The percentage of failed projects is high and it is despite the fact that every person involved in a project takes responsibility and understands the risks. According to the research conducted by Ernst & Young organization "Spotlight on oil and gas megaprojects" about 64% of all oil and gas projects have cost overruns and 73% of projects are behind the schedule. These indicators vary corresponding to region [2].

Most large projects involve complex structure and lots of contracts between several companies. In these situations, controlling the whole project state and making quick decisions can be challenging. Therefore project schedules can be delayed by slow responses from partners. It is also important because the situation in each part of project may have an affect on overall project execution time.

Project completion on time depends on rightly scheduled plan. In project management there is an approach Work Breakdown Structure, which decomposition of project into smaller tasks, aka activities. Each activity has its own duration time; it requires preconditions and provides the result. Activities are further used in project scheduling technique, which is chosen according the size, complexity and duration of the project. The technique must be practical and convenient. There are several methods that meet these criteria; two of

* Corresponding author e-mail: erke.shakenova@gmail.com

© 2016 NSP
Natural Sciences Publishing Cor.

them are Gantt charts and the critical path method (CPM). However Gantt charts have drawbacks in terms of displaying interdependency of activities, i.e. Gantt charts do not show how each activity depends on others. Moreover they require considerable time for updating [5].

Interdependent activities that made up the project build activity network. The complexity of the activity network in a large project increases rapidly with increase in the number of activities. CPM is a kind of network analysis method that is essential for making sense out of the mishmash of activities. Nevertheless one may not successfully apply CPM for large oil and gas projects. CPM uses only one estimated value for each activity duration time. However project activities are assumed to have uncertain durations, described by independent random variables with a known probability distribution function. In addition they are dynamic and their planned value may change under certain circumstances. Therefore Program Evaluation and Review Technique (PERT) was introduced.

The original PERT [8] is an activity-on-arrow network with one start and one finish event, which represent the beginning and the end of a project. To achieve the project, certain activities must be done according to a given pre-defined sequence. This logic is depicted by a directed, acyclic graph in which the vertices of the graph represent the events, while the arrows represent the tasks to be performed. An event happens when all preceding activities have been completed; only then one can start next tasks. In this way, the event is used for expressing logical dependencies between activities.

In a PERT network, activity durations are defined by stochastic variables that are supposed to be independent of each other. Generally, the distribution of the activity durations follows a so-called PERT-beta distribution. It uses subjective values *a*, *m* and *b*, commonly determined by a specialist, representing the optimistic (a), the most likely (m) and the pessimistic (b) durations of the activity.

However lots of research results show errors of this approach [9]. For example, the absence of theoretical justification and the unavoidable defects of the PERT statements were pointed out. Therefore many researchers suggest other distribution types for PERT estimations, like doubly truncated normal distribution (Kotiah and Wallace, 1973), the log-normal distribution (Mohan, et al., 2007), the mixed beta and uniform distribution (Hahn, 2008), the triangular distribution (Johnson, 1997) and the Parkinson distribution (Trietsch, et al., 2012)[4]. Nevertheless all the given distributions will fail if three-time estimations chosen wrongly. So, it is important to get more accurate values.

Several studies have investigated the case where activity times in a project are approximately known and more appropriately represented by fuzzy sets rather than crisp numbers. In specific, the problems of computing the intervals of possible values of the latest starting times and floats of activities with vague durations represented by fuzzy or interval numbers have fascinated intensively attentions and many solutions methods have been suggested [10].

However all this methods are used to construct networks for project activities. It is not possible to put interval or fuzzy values into simulation. So one cannot run simulations using results of suggested methods. Our approach is to get crisp numbers from results of survey taken from experts. It is done to further use those results in simulation.

2 Problem Statement

In practice, a question often arises as to how to obtain good estimates of a, m, and b. The process of defining these subjective values is called the PERT three-point estimation method. Several approaches can be used to get the required PERT three-point estimation. Some of the approaches are:

– Estimates defined by an experienced person
– Estimates extracted from standard time data
– Estimates obtained from historical data
– Estimates obtained from simple regression and/or forecasting
– Estimates generated by simulation
– Estimates derived from heuristic assumptions
– Estimates dictated by customer requirements

All of these approaches have disadvantages. For example, the drawback of using estimates furnished by an individual is that they may be inconsistent since they are limited by the experience of the person providing them. Individuals responsible for defining time estimates are usually not experts in estimation, and they generally have difficulty in providing accurate PERT time estimates, while asking them. The use of time standards, on the other hand, may not reflect the changes occurring in the current operating environment due to new technology, work simplification, new personnel, and so on. The use of historical data and forecasting is very popular because estimates can be verified and validated by actual records. In the case of regression and forecasting, there is the danger of extrapolation beyond the data range used for fitting the regression and forecasting models.

Commonly the values a, m and b in PERT analysis are subjectively determined by the person responsible for the completion of the activity. He is usually not a specialist in mathematical statistics, as is stated in some PERT studies, and to determine the most likely activity-time sometimes becomes a real problem for him. Their use can lead to large errors in the cumulative distribution function for project completion time whether or not multiple-critical-path aspects are handled correctly.

It has been claimed that fuzzy set theory is more appropriate to model uncertainty that is associated with parameters such as activity duration time, resource availability. In practice experts often estimate the values of project variables. Many of the values are defined based

© 2016 NSP
Natural Sciences Publishing Cor.

upon fuzzy and/or incomplete information. This type of information might be best modeled by fuzzy set techniques instead of probabilistic ones.

There are many concepts that use fuzzy logic as a base, for example when

– concept cannot be sharply defined (thus intrinsically fuzzy)
– notions are differently specified by each person
– no account is taken of the degree (but of linguistic labels).

Many approaches are taken in order to calculate critical path with fuzzy activity durations. So the problem is that fuzzy numbers taken from experts as a result of a survey are used to properly construct CPM charts. However one needs crisp number for each activity duration in order to be able to perform simulations. When Monte Carlo simulation is applied, the PERT distribution becomes effective tool for designing expert data. Moreover can be used to determine project risks and to define which resource's breakdown or activity's delay may lead to overall project's failure.

3 Preliminaries

In this section we review some technical terms and definitions used in fuzzy logic and project management.

Definition 1. *Let X be a space of points (objects), $X = \{x\}$. A fuzzy set F in X is characterized by a membership (characteristic) function $\mu F(x)$ which associates with each point in X a real number in the interval [0, 1], with the value of $\mu F(x)$ at representing the "grade of membership" of x in F [13]. This degree of membership is the primary formula of fuzzy set and there is no precise technique for finding the membership function, it is intuitive and practical.*

Definition 2. *The core of a fuzzy set $\widetilde{F}$ is a set C whose elements $x \in C$ have membership degrees $\mu F(x) = 1$.*

The specific types of fuzzy numbers are used as durations of activities in project management like bell, triangular, trapezoidal. For convenience in computation, a trapezoidal membership function is used in this paper for transformation of fuzzy number. Trapezoidal fuzzy numbers were chosen because the results of the survey is usually represented in four values.

Definition 3. *Trapezoidal Fuzzy Number $\widetilde{T}$ is defined as $\widetilde{T} = (a,b,c,d)$, where $a < b \leqslant c < d$ and $a > 0$. The parameters represent the following: lower limit a, an upper limit d, a lower support limit b, and an upper support limit c [11]. Basic relation between parameters is shown in Equation 1. It is also called number's shape function.*

$$\mu_T = \begin{cases} \frac{x-a}{b-a} & \text{when } x \in [a,b) \\ 1 & \text{when } x \in [b,c] \\ \frac{d-x}{d-c} & \text{when } x \in (c,d] \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

The most often used operators for fuzzy PERT/CPM are addition, subtraction, maximum, minimum and ranking. Addition operator is applied to calculate the earliest completing times and overall project completing time. Subtraction operator is used to compute the latest starting and completing times. Maximum operator is applied for earliest starting times and minimum operator is used in calculating latest starting times. Suppose we have two triangular fuzzy numbers are defined $\widetilde{T_1} = (a_1,b_1,c_1,d_1)$ and $\widetilde{T_2} = (a_2,b_2,c_2,d_2)$. The most often used formulas of addition, subtraction, maximum and minimum are the following [12]:

Addition: $\widetilde{T_1} + \widetilde{T_2} = (a_1+a_2, b_1+b_2, c_1+c_2, d_1+d_2)$
Subtraction: $\widetilde{T_1} - \widetilde{T_2} = (a_1-d_2, b_1-c_2, c_1-b_2, d_1-a_2)$
Maximum: $max\{\widetilde{T_1},\widetilde{T_2}\} = (max\{a_1,a_2\}, max\{b_1,b_2\}, max\{c_1,c_2\}, max\{d_1,d_2\})$
Minimum: $min\{\widetilde{T_1},\widetilde{T_2}\} = (min\{a_1,a_2\}, min\{b_1,b_2\}, min\{c_1,c_2\}, min\{d_1,d_2\})$

Ranking fuzzy numbers applied in fuzzy PERT is used to determine the earliest starting time. Techniques for ranking fuzzy numbers are abundant in the literature. According to the investigation by X.Wang and E.Kerre this problem of ordering fuzzy quantities has been analyzed by many researchers [14]. If $a_1 \leqslant a_2, b_1 \leqslant b_2, c_1 \leqslant c_2$ and $d_1 \leqslant d_2$ the ranking of $\widetilde{T_1}$ and $\widetilde{T_2}$ is said that $\widetilde{T_2}$ is strongly greater than $\widetilde{T_1}$. If one of these four inequalities is not true, the comparison rule has to take the advantage of weak comparison rule (WCR).

Moreover in order to represent to the user not intervals of fuzzy numbers, but crisp numbers we should use defuzzification method. This is the most important part of the process. In paper W.Leekwijck and E.Kerre different defuzzification techniques are explained properly and shown which type can satisfy what kind of criteria [15]. In this paper we will use First of maxima (FOM) method. It like other maxima methods chooses one element from the core of given fuzzy set. This criterion is called semantically correct defuzzification and it is computationally very efficient.

$$FOM(\widetilde{F}) = min\ core(\widetilde{F}) \quad (2)$$

In order to compute critical path for activities with fuzzy time durations our system uses following steps:

1. Compute the earliest starting fuzzy time for each activity
2. Calculating the earliest completion/finishing times for each activity
3. Computing the latest starting fuzzy time for activities
4. Calculate the latest starting times of all activities

© 2016 NSP
Natural Sciences Publishing Cor.

5. Compute slack time, which is also fuzzy by one of the following formulas

Definition 4. *Activities with no slack are defined as critical.*

Different formulas are used to calculate mean and variance according to the shape of distribution [6]. In this paper we applied most widely used rules for beta distribution [7]. Using three-time estimates in PERT formulas:

$$t_e = \frac{a+4m+b}{6} \tag{3}$$

$$s^2 = \frac{(b-a)^2}{36} \tag{4}$$

Equation 3 represents expected time for activity duration and Equation 4 is a formula for variance of the duration.

4 Solution

Our goal is to extract crisp values from given fuzzy expert survey results. The solution is developed as web application. Since experts may be located in different places, they can fill in the questionnaire from any place. Only presence of Internet connection is required.

All values for time estimates will be stored in database. MySQL database management system is used to store all the data. The interface developed allows experts to answer questions in natural language and to obtain useful results, without having to modify neither the structure of the database nor the DBMS query language and without knowing database table structures. In Figure 1 there represented structure of database tables.

For activities we store three values for duration (optimistic, most likely and pessimistic) and actual duration of already completed projects. Moreover table contains columns act_id - activity identification number, project_id - ID of project of given activity, act_title, act_desc - description and title of activity, is_critical - whether or not activity is on critical path.

Each expert will be given a username and password, using them he can login into system. In main page all available projects will be shown. In this page expert chooses one project and can see its activity list and precedence network, i.e. see the relationships between activities. After that expert may add new answer to questionnaire or change previously added value. Main distinction of our method from other methods is that experts will choose not exact numbers, but intervals for each given fuzzy number. So for each activity experts will be asked to indicate interval that can be defined as short, medium (normal) and long period for given activity. Since different activities' durations might differ in measurements, there also will be choice for appropriate time term. For example, some activities can be performed during some hours, while others' execution might take several weeks or months.

Another important note about output of this survey. As an output one should get three time estimates for each activity: optimistic, most-likely and pessimistic. Asking from one expert all three values will be too much and

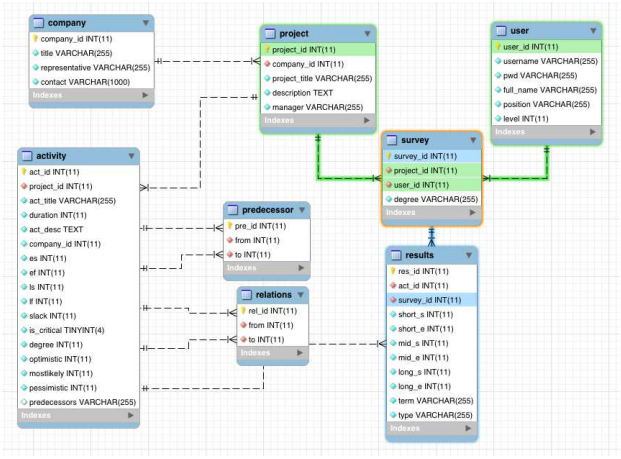

Fig. 1: Database structure.

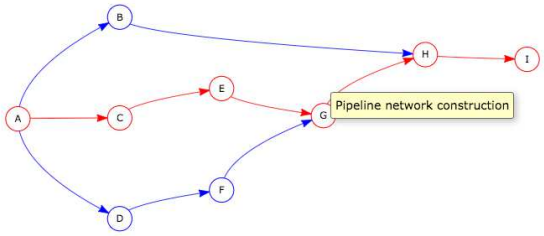

Fig. 2: Network of activities with critical path.

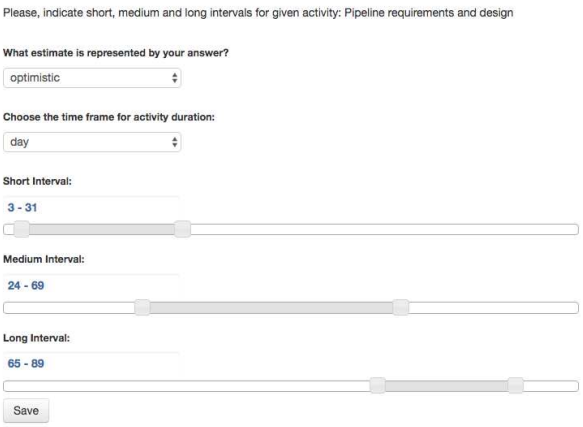

Fig. 3: Example of the question for expert.

© 2016 NSP
Natural Sciences Publishing Cor.

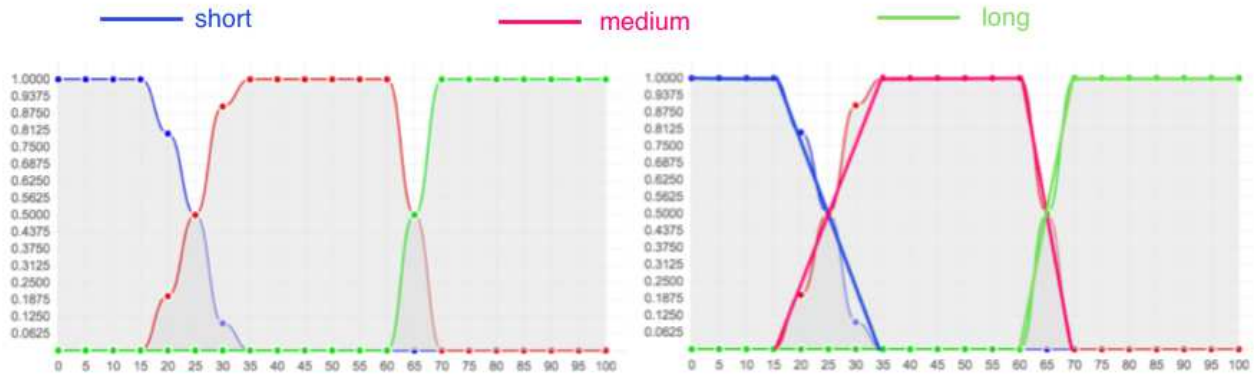

Fig. 4: Approximation of membership functions of activity durations.

inconvenient, therefore we ask experts to indicate for which estimate their answer is most suitable. This can be considered as another intrinsic for human nature approach like fuzzy logic. Because some people might describe themselves as pessimists or inexperienced person might be too optimistic because of lack of failure situations in experience.

As the result of survey we will get fuzzy numbers for each activity durations' three estimates. The final step will be defuzzification of given values. Then this result can be used as an input for simulation using beta-PERT distribution.

5 Demonstration of solution

In this paper we consider one of the examples of Oil and Gas projects. It is the example of Pipeline construction project. Activities are represented in the Table 1.

Table 1: Activities of example project.

#	Title	Activity
1	A	Pipeline requirements and design
2	B	Simulation and optimization
3	C	Land pipelines' selection and procurement
4	D	Offshore pipelines' selection and procurement
5	E	Land pipelines' transportation
6	F	Offshore pipelines' transportation
7	G	Pipeline network construction
8	H	Pipeline welding and lowering
9	I	Testing and finishing

When expert selects one project list of its activities and activity-on-node network will be represented. After completion of all the surveys it will possible to construct critical path for the project using most-likely values as activity durations. Critical path on project's page in our web application is a graph in Figure 2. Red circles represent activities in critical path and blue circles represent activities with slack, which is strictly greater than zero. However at the beginning all the activities will be blue, because their duration is not known yet. For the ease to read purpose on network we show only short titles (letters) of activities, but full title will appear when user hover with mouse to the short title.

All the answers taken from experts about activity duration will be used to build fuzzy sets for activity durations. This method can be used when project manager doesn't know exact ranges for activity durations. Then after getting all the values project manager can build fuzzy sets for each linguistic variable (short, normal, long) by approximation of obtained values. Example of approximation of membership functions is shown in Figure 4.

After collecting information from all experts defuzzification method described above will be used to get crisp number - actual percentage of effect. In other words we are interpreting the membership degrees of the fuzzy sets into a specific decision or real value. We will get core of the fuzzy number. $core(\widetilde{F}) = \{35, 40, 45, 50, 55, 60\}$. Using FOM method, we can find that answer will be 35 days.

6 Conclusion

Our proposed model uses fuzzy sets and linguistic variable in a survey for experts. Questions will be generated for each activity in a project. We also use fuzzy numbers for activity durations, since they can't be clearly defined. There are lots of researches on calculation of critical path activities; therefore we used most popular ones. Contribution of paper is survey method is used to determine risks in large Oil and Gas projects. This

© 2016 NSP
Natural Sciences Publishing Cor.

method of survey is useful for PM because of several reasons:

– since answers are intervals more qualitative inputs will be given to simulation
– survey will be taken from different experts, so result is not subjective opinion of one person
– survey will be taken from different experts, so result is not subjective opinion of one person
– experts can participate in polling remotely, no need to gather experts in place, which is time consuming operation
– questions are linguistic variables, which are more close to human nature
– Oil and Gas projects usually large sized, as a result involve lots of risks and require experts' knowledge.

As a further development different types of question can be added to the survey. For example, questions about resources, costs and other parameters of the project. Moreover one might investigate and reveal advantages and disadvantages of using different kinds of defuzzification methods.

References

[1] A Guide to the Project Management Body of Knowledge (PMBOK Guide) Fifth Edition. PMI (2013).
[2] Preiss A. Spotlight on Oil and Gas Megaprojects. EY (2014).
[3] McKenna M.G., Wilczynski H., VanderSchee D. Capital Project Execution in the Oil and Gas Industry. Booz Allen Hamilton Inc (2006).
[4] Dan Trietsch, Kenneth R. Baker. PERT 21: Fitting PERT/CPM for use in the 21st century. International Journal of Project Management **30**, 490-502 (2012).
[5] Adedeji B. Badiru, Samuel O. Osisanya. Project Management for the Oil and Gas Industry. A World System Approach. CRC Group (2013).
[6] Joseph J.Moder, E.G.Rodgers. Judgment Estimates of the Moments of PERT Type Distributions. Management Science **15**, No. 2, B76-B83 (1968).
[7] Jamal A. Al-Saleh, Satish K. Agarwal. Extended Beta Distribution and Mixture Distributions with applications to Bayesian analysis. Journal of Statistics Applications & Probability, NSP **2**, No. 1, 61-72 (2013).
[8] Dimitri Golenko-Ginzburg. On the Distribution of Activity Time in PERT. The Journal of the Operational Research Society **39**, No. 8, 767-771 (1988).
[9] Miklos Hajdua, Orsolya Bokora. The Effects of Different Activity Distributions on Project Duration in PERT Networks. Procedia - Social and Behavioral Sciences **119**, 766-775 (2014).
[10] P. Zielinski. On computing the latest starting times and floats of activities in a network with imprecise durations. Fuzzy Sets and Systems **150**, 53-76 (2005).
[11] Salim Rezvani, Mohammad Molani. Representation of trapezoidal fuzzy numbers with shape function. Annals of Fuzzy Mathematics and Informatics **8**, No. 1, 89-112 (2014).
[12] Daniyar Artykov, Lyazzat Atymtayeva. A Fuzzy Linear Programming Approach for Resource-Constrained Project Scheduling. Advanced Engineering Technology and Application **4**, No. 3, 47-52 (2015).
[13] L.A. Zadeh. Fuzzy Sets. Information and Control **8**, 338-353 (1965)
[14] Xuzhu Wang, Etienne E. Kerre. Reasonable properties for the ordering of fuzzy quantities (I). Fuzzy Sets and Systems **118**, 375-385 (2001).
[15] Werner Van Leekwijck, Etienne E. Kerre. Defuzzification: criteria and classification. Fuzzy Sets and Systems **108**, 159-178 (1999).

Yerkezhan Shakenova received the bachelor degree in Information Systems at Kazakh-British Technical University in Almaty. Her main research interests are: fuzzy logic, project management, reliability management, Oil and Gas projects, web application development.

© 2016 NSP
Natural Sciences Publishing Cor.