Planning and Control of Projects with Different Types of Precedence Relationships and a Service Level Using Stochastic Simulation

D.F. Muñoz
Department of Industrial & Operations Engineering
Instituto Tecnológico Autónomo de México
México DF, México

D.F. Muñoz
Department of Biomedical Informatics
Stanford University
USA

***Abstract*—We present a decision support system (DSS) to compute performance measures of a project under uncertainty on the activities' duration as well as four different types of precedence relationships. The DSS generates replicates of the project's performance, in which we simulate the duration of each activity. From these replicates, the expected completion time, the variance of completion time, the service time for a given service level and the probability that each activity will be in the critical path are estimated along with their corresponding measures of accuracy. A validation of the DSS was performed by computing the empirical coverage, mean and standard deviation of half-widths, mean square error and empirical bias for the main performance metrics of a given project. Our experimental results show that the procedures implemented in the DSS provide a good coverage and consistent half-widths.**

Keywords-Project Management; PERT; CPM; Project Simulation

I. INTRODUCTION

A project is a process that requires the execution of multiple activities with different types of dependency relationships. These activities are tasks associated with the execution of the project and consume a certain amount of time to be completed. In turn, activities may also have precedence relationships amongst them; in other words, the initiation (or completion) of specific tasks may require that other activities have been completed (or initiated). There are several project management techniques that can be used to estimate the performance of a project, all of which require the identification of the activities associated to the project, their durations and precedence relationships.

The most commonly used performance measures of a project are: the total duration (time between initiation of the first activity and completion of the last activity) and its cost; both of which are typically dependent on the durations of the individual project's tasks. As is known from the project management literature, the classic methodology to estimate the duration and cost of a project is based on the use of the PERT/CPM approach [1], which has been used since the 1950s. The critical path method (CPM) is used to estimate the total duration of the project when the duration of the activities is known. On the other hand, classical program evaluation and review technique (PERT) incorporates uncertainty in the duration of these activities, but relies on the expected durations to calculate the critical path. Thus, the validity of the PERT/CPM approach relies on several assumptions, such as: the duration of the activities are statistically independent, all precedence relationships must be of finish to start type (i.e., precedent activities must have concluded for the dependent activity to start) and there is only one possible critical path, which is determined from the expected duration of the activities. In reality, it is quite difficult for a project to meet all these requirements since there may be several types of precedence relationships aside from the traditional finish to start (e.g., start to start) and the uncertainty in the duration of individual activities will generate uncertainty in the critical path. Due to these limitations, several other techniques for project performance analysis have become widespread and rely on more robust methods, such as stochastic simulation [2].

The uncertainty associated to the duration of individual activities is incorporated into project analysis using probability distributions. However, due to the precedence relationships across different activities, it is not possible to obtain analytical expressions for the performance measures estimates; but they may be estimated through simulation. The use of stochastic simulation to estimate performance measures of a project has been widely used in the project management community. For instance, a method to calculate the critical path of a project using the classical PERT/CPM approach and stochastic simulation is proposed in [3]. A software (called SPSS) that is capable of estimating the probability that the duration of a project does not exceed a total (user defined) time is developed in [4]. In subsequent work [5], an update to the SPSS software (named S3) is reported; S3 is capable of allowing the user to define an accuracy level for the estimates by providing confidence intervals, as well as the number of replicates per simulation experiment. All these models were constructed assuming only one type of precedence relationship: the classic finish to start. A simulation-based decision support system (DSS) to estimate the expected duration of a project under different types of precedence relationships is developed in [6] and, in this work, we extend the use of simulation to estimate risk measurements such as the variance of the project duration and the service time for a given service level.

It is relevant to note that the methods described here belong to class of algorithms that leverage on the use of simulation to solve decision-making problems in project management. Other examples of work in this area include a

© 2015. The authors - Published by Atlantis Press

report on the use of stochastic simulation to determine the crash times that minimize the cost of different projects [7], the use simulation to determine optimal resource allocation across several projects in the context of construction industry [8], and the use a simulation approach to determine optimal resource allocation for the activities of a specific construction project [9, 10]. Once again, in all this prior work, only the classical finish to start precedence relationship was assumed.

II. DSS Features

Input data for our DSS consists of a detailed list of project tasks, including duration and precedence relationships for each. We include the following types of precedence relationships [11]: finish to start (FS, the activity may start only if the preceding activities have concluded), start to start (SS, the activity may start only if the preceding activities have also started), finish to finish (FF, the activity may finish only if the preceding activities have also concluded) and start to finish (SF, the activity may finish only if the preceding activities have started).

The classical approach to calculate the critical path of a project, as described with the CPM method [1, 12], relies on an algorithm that calculates the early start time (EST), early finish time (EFT), late start time (LST) and late finish time (LFT) for each activity. The algorithm runs in two steps: a "forward" and "backward" iteration. In the "forward" iteration, the EST and EFT (for each activity) are calculated from their predecessors, starting with the first activities (those with no predecessors). On the other hand, in the "backward" iteration, the LST and LFT (for each activity) are calculated from their successors, starting with the last activity (those completed at the end of the project). Having made these calculations, activities with EST equal to their LST (and, in consequence, also have equal EFT and LFT) belong to the critical path, since there is no slack between the early and late starting times. It is worth mentioning that this algorithm is designed for projects that work with FS precedencies only.

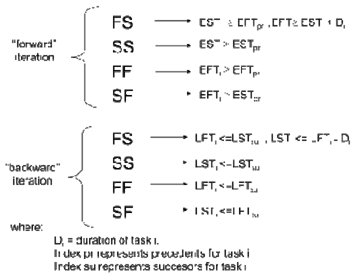

FIGURE I. .COMPUTATIONS REQUIRED TO DETERMINE EST, EFT, LST AND LFT.

We developed a variant of the CPM to identify the critical path of a project with more than one type of precedence relationship. Said algorithm is also divided into two steps: a "forward" and "backward" iteration, where the EST, EFT, LST and LFT of each activity are sequentially calculated. However, contrary to the traditional method, several additional conditions for each type of precedence relationship must be verified to determine the critical path. In Figure. 1, we show the required calculations (per iteration of the algorithm) according to the precedence relationship that must be satisfied.

Note that the EST and the EFT are the lower bounds for the inequalities and the LST and LFT are the upper bounds. On the other hand, we assume that the network of activities that form of a project do not contain cycles, making it possible to sort them in sequence for each of the algorithm's iterations.

For our DSS, we used triangular distributions to model activity durations. The three-parameter duration values for these distributions, as defined by PERT, were: optimistic, probable and pessimistic. The DSS uses Monte Carlo simulation to generate replicates of the project's performance. In other words, in each replication, we simulate the duration of every activity, obtain the critical path and calculate the total duration of the project. Therefore, for a number of replications, we calculate the mean and variance of project durations, and estimate the probability that each activity will be part of the critical path, along with the corresponding accuracy measures (half-widths of the confidence intervals, see [13]). Figure. 2 shows the pseudo-code used in the DSS to obtain the durations D_i , $i=1,\ldots,m$ (where m is the number of replications), from which the point estimators for the expectation and variance of project's duration, service time and probability that a task (i) be in the critical path, respectively, are as follows.

$$\hat{\mu}_D = \frac{1}{m}\sum_{i=1}^{m} D_i , \quad \hat{\sigma}_D^2 = \frac{1}{m}\sum_{i=1}^{m}(D_i - \hat{\mu}_D)^2 , \quad \hat{T}_\alpha = D_{(\lceil m\alpha \rceil)} , \quad \hat{p}_j = \frac{1}{m}\sum_{i=1}^{m} I_{ij} , \quad (1)$$

where $0<\alpha<1$ is a given service level, $D_{(1)} \le D_{(2)} \le \ldots \le D_{(m)}$ denote the ordered D_i 's, and I_{ij} is 1 when the *j*-th task was in the critical path in the *i*-th replication, and 0 otherwise.

```
Input data for the project network.
For i = 1 to the number of replications m
    Simulate task durations.
    Apply CPM and compute the project duration D_i.
End for
Compute point estimators and corresponding halfwidths
```

FIGURE II. PSEUDO-CODE FOR THE DEVELOPED DSS.

As discussed in most textbooks on stochastic simulation, the number of replications m must be large enough to ensure that the point estimators defined in (1) fall within a given accuracy of the corresponding parameter, and the most commonly used measure of accuracy in the stochastic simulation literature is the half-width of a $100(1-\beta)\%$ asymptotic confidence interval for the corresponding parameter. For the estimators provided in (1), the corresponding $100(1-\beta)\%$ half-widths are given by

$$H_{\hat{\mu}_D} = z_\beta \hat{\sigma}_D / \sqrt{m} , \quad H_{\hat{\sigma}_D^2} = z_\beta \hat{\sigma}_{\hat{\sigma}} / \sqrt{m} , \quad H_{\hat{T}_\alpha} = \left(D_{(n_2)} - D_{(n_1)}\right)/2 ,$$
$$H_{\hat{p}_j} = z_\beta \hat{\sigma}_j / \sqrt{m} , \quad (2)$$

where $0<\beta<1$, z_β is the $(1-\beta/2)$ quantile of a standard normal distribution, $n_1 = \left\lceil m\alpha - z_\beta [m\alpha(1-\alpha)]^{1/2} \right\rceil$, $n_2 = \left\lceil m\alpha + z_\beta [m\alpha(1-\alpha)]^{1/2} \right\rceil$, $\hat{\sigma}_j^2 = \left[\sum_{i=1}^{m}(I_{ij} - \hat{p}_j)^2\right]/(m-1)$, and

$\hat{\sigma}^2_{\hat{\sigma}} = 4\bar{D}_1^2\hat{\sigma}_D^2 - 4\bar{D}_1 S_{12} + S_{22}$, where $\bar{D}_i = m^{-1}\left(\sum_{k=1}^{m} D_k^i\right)$, $i = 1,2,3,4$, $S_{12} = \bar{D}_3 - \bar{D}_1\bar{D}_2$, $S_{22} = \bar{D}_4 - \bar{D}_2^2$. The asymptotic validity of the half-widths defined in (2) rely on a Central Limit Theorem for each of the point estimators defined in (1) and, in particular, the asymptotic validity of the corresponding confidence interval for the service time is provided in [14]. It is worth mentioning that a test version of our DSS (using an Excel interface) can be downloaded at http://ciep.itam.mx/~davidm/sofdop.htm.

III. PERFORMANCE VALIDATION FOR THE DSS

As shown in Figure. 2, when a parameter is estimated using stochastic simulation, the experiment is repeated a number of times to improve the accuracy of the point estimators. In addition, the estimation procedures must be consistent, in the sense that the estimators must approach the parameter values as the number of experiment replications is increased. To empirically verify that the methodologies implemented in our DSS provide consistent estimators, we tested the system using a hypothetical project with 18 tasks (see [15] for details) and repeated the estimation procedure (Figure. 2) M times for different values of m. Thus, it was possible to calculate measures that quantify the performance of the DSS such as: empirical coverage (EC), mean and standard deviation of the half-widths, mean squared error (MSE) and the empirical bias (B). The calculations are detailed as follows.

Let h_i be the half-width obtained in experiment i, for $i = 1,\ldots,M$, the mean and standard deviation of the half-widths are the mean and standard deviation calculated using all h_i, respectively. On the other hand, if $\hat{r}_i$ is the estimator for parameter r obtained from experiment i, for $i = 1,2,\ldots,M$, the mean square error (MSE) is defined by:

$$MSE = \frac{\sum_{i=1}^{M}(\hat{r}_i - r)^2}{M} \tag{3}$$

using the same notation, the empirical coverage is defined as:

$$EC = \frac{1}{M}\sum_{i=1}^{M} C_i \tag{4}$$

where $C_i = 1$ if $|\hat{r}_i - r| < h_i$, and $C_i = 0$ otherwise; the empirical bias is:

$$B = \frac{1}{M}\left(\sum_{i=1}^{M}\hat{r}_i\right) - r \tag{5}$$

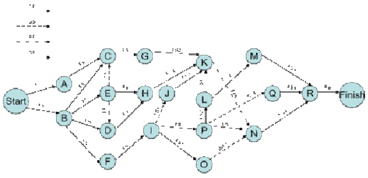

FIGURE III. ACTIVITIES DIAGRAM FOR THE HYPOTHETICAL PROJECT

It is relevant to note that, if the DSS performs adequately, the MSE, the empirical bias and the mean of the half-widths should converge to zero as the number of replications increases; and in this case, the empirical coverage will also converge to the nominal value of $(1-\beta)$.

Figure. 3 shows the flow diagram for the project used to perform the experiments. Note that a number on an arc indicates the number of days that must elapse for the activity to start following the fulfillment of the precedence condition. If there is no number, we assume a value of zero, i.e., that the activity may start immediately after the precedence has been met.

IV. RESULTS

In a first experiment we run $m = 1.6E07$ simulation replications in order to obtain very accurate estimates. From this experiment we obtained an expected duration of $\hat{\mu}_D = 44.525$, with a 95% half-width of 0.001, variance of duration was $\hat{\sigma}_D^2 = 5.035$ with a 95% half-width of 0.003, service time for a 90% service level was $\hat{T}_{0.9} = 47.441$ with a 95% half-width of 0.002, and similarly we obtained estimates of the probabilities of being in the critical path for every activity, e.g., $\hat{p}_{16} = 0.7493$ with a 95% half-width of 0.0002.

The point estimators obtained in our first experiment were considered as the true parameter values and we repeated the estimation experiments   times with a 95% confidence level. As a proof of concept, we will only show the results obtained for the probability of being in the critical path of a single task *i* = 16. Tables 1, 2, 3 and 4 summarize the results found using different values of m. Note that even for relatively small values of m, we find coverage very close to the nominal value (0.95), as well as half-widths that decrease as m increases. For the case of variance estimation, we obtained over-coverage which may be explained from the fact that the corresponding variance estimation is biased.

TABLE I .RESULTS FOR THE MEAN DURATION OF THE PROJECT FOR $M = 1000$.

		95% Half-width			
	EC	**Mean**	**Std. Dev.**	**MSE**	**Bias**
m = 400	0.936	0.2201	0.0075	0.0136	-0.0035
m= 1600	0.938	0.1100	0.0019	0.0034	-0.0038
m= 4800	0.955	0.0635	0.0006	0.0010	-4.1E-06

TABLE II .RESULTS FOR THE DURATION VARIANCE OF THE PROJECT FOR $M=1000$.

	EC	95% Half-width Mean	Std. Dev.	MSE	Bias
m=400	1	1.5388	0.0624	0.1179	0.0162
m=1600	0.994	0.4806	0.0132	0.0301	0.0030
m=4800	0.98	0.2245	0.0042	0.0091	0.0007

TABLE III .RESULTS FOR THE 90% SERVICE TIME OF THE PROJECT FOR $M=100$.

	EC	95% Half-width Mean	Std. Dev.	MSE	Bias
m = 400	0.952	0.3771	0.0750	0.0387	-0.0178
m=1600	0.946	0.1896	0.0271	0.0096	-0.0082
m = 4800	0.954	0.1086	0.0118	0.0030	-0.0022

TABLE IV .RESULTS FOR THE PROBABILITY OF TASK 16 IN CRITICAL PATH FOR $M = 1000$.

	EC	95% Half-width Mean	Std. Dev.	MSE	Bias
m = 400	0.957	0.0424	0.0012	0.0004	0.0014
m= 1600	0.945	0.0212	0.0003	0.0001	-0.0001
m = 4800	0.951	0.0123	0.0001	4.1E-05	0.0001

V. CONCLUSIONS

In this work, we successfully developed a DSS capable of estimating performance measures associated to the total duration of a project that incorporate the uncertainty in the duration of individual activities and more than one type of precedence relationship between these. We used stochastic simulation to incorporate the uncertainty in the activities durations and we developed an algorithm capable of determining the critical path for each project replication. In turn, this allowed us to estimate the expected value for the total duration of the project, the variance of completion time, the service time for a given service level, and the probability that each activity will belong to the critical path and compute error measures for each of these by providing half-widths of a confidence interval.

Lastly, we used a hypothetical example to test the performance of the developed DSS by measuring the accuracy and consistency of the reported estimates. The results from this hypothetical test show that even for a small number of replications, the system is capable of finding good empirical coverage with half-widths that decrease as the number of repetitions increase.

ACKNOWLEDGMENTS

We thank the Asociación Mexicana de Cultura A.C. for their support in the development of this work.

REFERENCES

[1] Hillier, F.S. & Hillier, M.S., Introduction to Management Science: A Modeling and Case Studies Approach with Spreadsheets (5th ed.), McGraw-Hill: New York, 2014.

[2] Meredith, J.R. & Mantel, S.J., Project Management: A Managerial Approach (8th ed.), John Wiley: New Jersey, 2012.

[3] Lu, M. & AbouRizk, S.M., Simplified CPM/PERT simulation model. Journal of Construction Engineering and Management, 126(3), pp. 219-226, 2000.

[4] Lee, D., Probability of project completion using stochastic project scheduling simulation. Journal of Construction Engineering and Management, 131(3), pp. 310-318, 2005.

[5] Lee, D. & Arditi, D., Automated statistical analysis in stochastic project scheduling simulation. Journal of Construction Engineering and Management, 132(3), pp. 268-277, 2006.

[6] Muñoz, D.F. & Muñoz, D.F., Planeación y control de proyectos con diferentes tipos de precedencias utilizando simulación estocástica. Información Tecnológica, 21(4), pp. 25-33, 2010.

[7] Kuhl, M.E. & Tolentino-Peña, R.A., A dynamic crashing method for project management using simulation-based optimization. Proc. of the 2008 Winter Simulation Conference, eds. S.J. Mason, R.R. Hill, L. Mönch, O. Rose, T. Jefferson & J. W. Fowler, IEEE: Piscataway, pp. 2370-2376, 2008.

[8] Hoi-Ching L. & Ming, L., Simulation-based optimized scheduling of limited bar-benders over multiple building sites. Proc. of the 2008 Winter Simulation Conference, eds. S.J. Mason, R.R. Hill, L. Mönch, O. Rose, T. Jefferson & J. W. Fowler, IEEE: Piscataway, pp. 2353-2360, 2008.

[9] Liu, Y. & Mohamed, Y., Multi-agent resource allocation (MARA) for modeling construction processes. Proc. of the 2008 Winter Simulation Conference, eds. S.J. Mason, R.R. Hill, L. Mönch, O. Rose, T. Jefferson & J. W. Fowler, IEEE: Piscataway, pp. 2361-2369, 2008.

[10] Tang, P., Mukherjee A. & Onder, N., Construction schedule simulation for improved project planning: activity criticality index assessment. Proc. of the 2013 Winter Simulation Conference, eds. R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill & M. E. Kuhl, IEEE: Piscataway, pp. 3237-3248, 2013.

[11] Chatfield, C. & Johnson, T., Microsoft Office Project 2007 Step by Step, Microsoft Press: Redmond, 2007.

[12] Muñoz, D.F., Administración de Operaciones: Enfoque de Administración de Procesos de Negocios, Cengage: Mexico City, 2009.

[13] Ross, S.M., Simulation (5th ed.), Prentice Hall: New Jersey, 2010.

[14] Serfling, R.J., Approximation Theorems of Mathematical Statistics, John Wiley: New York, 1980.

[15] Torres, K.G., Simulación de un Proyecto con Diferentes Relaciones de Precedencia. Industrial Engineering Dissertation, Instituto Tecnológico Autónomo de México, Mexico City, 2008.