RESOURCE ALLOCATION IN STOCHASTIC PROJECT NETWORKS

John M. Burt, Jr.
Graduate School of Management, U.C.L.A.
Los Angeles, California 90024

SUMMARY

This paper is concerned with the allocation of a single type of resource in project networks (such as the PERT representation) composed of activities whose processing times are random variables. There is a continuum of alternative processing methods for each activity and corresponding costs (resource utilizations) associated with each method. The effects of these different processing methods for any given activity are represented by modifications in the probability functions of the time required to complete that activity. Over the duration of a project, the decision maker (i.e., project manager) must decide which method (i.e., probability function) should be used for processing each activity. This sequential decision making must be performed so as to minimize the expected completion time of the entire project in light of restrictions on overall resource availability. A simulation study was performed to analyze the comparative effectiveness of several decision rules for "solving" this problem. This paper presents the results of the initial stages of that study.

INTRODUCTION

Over the past decade, network graphs have received considerable attention as aids in the planning and control of large-scale projects. There is an abundance (plethora may be a more apt word) of articles in the professional journals presenting various problem formulations and techniques for analyzing these network graphs. It has been recently noted that the results obtainable via network analysis are seldom utilized, or sometimes even considered, by decision makers in top level management; see the editorial by Vazsonyi.[9] In the views of this author, this behavior of management may be attributed, in large part, to the failure of existing network analysis models to deal simultaneously with the following three realities of actual project management:

(1) Uncertainty; the time required to complete an activity is seldom known in advance with certainty. This is particularly the case for typically non-repetitive "once-in-a-lifetime" projects.

(2) Resources; the time required to complete most activities is seldom independent of the resources allocated to them. The allocation of limited resources amongst "competing" activities is a primary function of the project manager.

(3) Decision making is not a static concept: Over the duration of a project, the project manager sequentially re-allocates resources in light of new information on the status of the project. Resource allocation in projects is seldom executed in the immutable manner of a "budgeting" process.

While none of the research literature mentioned above deals simultaneously with all three of these factors, considerable attention has been given to problems involving the factors singularly. Even a partial listing of appropriate references would be voluminous; the interested reader may consult the excellent bibliographies in Wiest and Levy,[10], or in Elmaghraby.[5]

This paper presents empirical results from a project management model which incorporates the above three aspects of the problem. The comparative effectiveness of several resource allocation rules is analyzed in an environment of uncertain activity times and sequential decision making. Due to the complexity and stochastic nature of the problem, the use of simulation was essential to the analysis. In the remaining sections, the particular problem under study is formulated, followed by a short discussion of the different allocation decision rules. Simulation results are then presented and interpreted. Finally, since this paper represents only the first stage of this study, the directions of future research and some associated difficulties in experimental design are discussed.

PROBLEM FORMULATION

The model under consideration is one in which a limited resource of a single type(1) is allocated amongst the activities in a specific project. The objective of this allocation decision making is to minimize expected project completion time, i.e., the time required to finish the entire project on the average. The time required to complete each activity is uncertain, and hence, is denoted by a random variable. The effect of allocating a particular level of resource to an activity is represented by the probability function of the time to complete that activity, given that resource level. There is a cost (resource utilization) associated with the amount of resource given to any activity.

---

(1) The only type of resource being considered in this paper is one that is used up in the processing of an activity and hence cannot be reallocated to a second activity later on in the project's life. Thus, money or raw materials would be considered a resource in this sense, but a machine or laborer would not.

Two major types of problems will be considered; they correspond to allocations which are made statically (i.e., only once, prior to the start of the project) and those which are made sequencially (i.e., repetitively, over the duration of the project's life). To formalize these concepts, some notation will be introduced and the problem will be represented in symbolic form.

Let: The network be that of Figure 1, where events (nodes) are denoted by the circles, and activities-precedence relations (directed arcs) are denoted by the arrows.

- X_i be the level of resource allocated to activity i, where $0 \leq X_i \leq 1$. X_i is the decision variable.
- R_i be a uniformly distributed random number over the interval (0,1) associated with the time required to complete activity i.
- $200(1-X_i)R_i$ be the time required to complete activity i, given that X_i amount of resource is allocated to it. Thus, the times required to complete activities are independent random variables uniformly distributed over the interval $(0, 200(1-X_i))$.
- X_i^2 be the cost of allocating X_i units of resource to activity i. These quadratic costs reflect the effect of "diminishing returns". Thus, the cost of (probabalistically) reducing the time required to complete activity i by a given amount increases as X_i varies from 0 (i.e., no allocation) to 1 (i.e., the maximum allocation).
- $\{j\}$ be the set of indices of activities lying on path j.

The static problem is that of allocating the resource amongst activities so as to minimize expected project completion time, subject to a limitation on resource availability[2]. That is,

P1: Determine X_i, i = 1,2,...,15, so as to minimize expected project completion time,

$$= E\left\{ \max_{j=1,2,\ldots,10} \left[\sum_{i \in \{j\}} 200(1-X_i)\, R_i \right] \right\}$$

Subject to the "budget" constraint,

$$\sum_{i=1}^{15} X_i^2 \leq 2$$

The sequential problem is similar in form; however it is extremely difficult to write down concisely. The difficulty lies in the fact that over the course of the project's duration, the X_i's become fixed values rather than decision variables (i.e., decisions have been made and carried out) and the R_i's become fixed rather than random variables (i.e. activity times "early" in the project become known). Thus in the sequential problem, the X_i's are "fixed" at different points in time in an order that depends upon the changing set of previously fixed X_i's and their associated R_i's.

It is computationally arduous, if not impossible, to solve even the static problem, P1, with analytic methods. By conditioning on random variables associated with those activities lying on more than one path, it would be possible to write down an open-form expression for the expected project completion times as a function of the X_i'x. The open-form expression would involve sixteen integrals---one for each random variable contributing to project completion time plus an additional integral for the expectation. Deriving a closed-form solution, which could then be optimized with

---

(2) If the available amount of resource (the budget) is too large, then most X_i's will be 1 (i.e. most activities would receive a maximum allocation) and if the budget is too low, then most X_i's will be 0. Sensitivity analysis showed that a budget equal to 2 avoided these extreme "solutions" and hence that value was used in all the empirical studies.

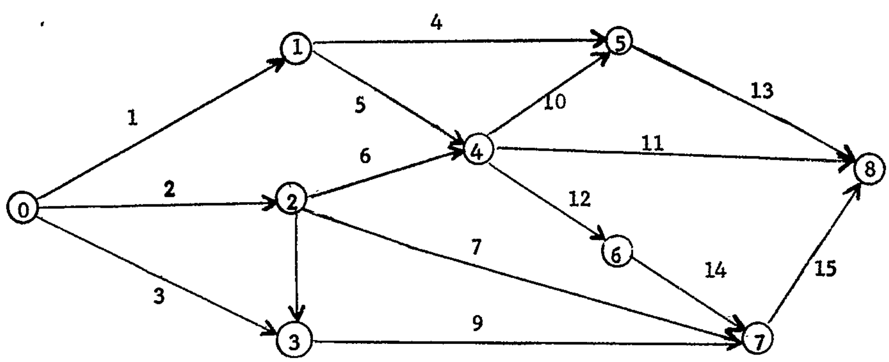

Figure I : Network Under Consideration

respect to the X_i's would be extremely difficult even if approximate methods (e.g., numerical integration) were employed. To obtain a deeper understanding of the use of analytic methods in stochastic network analysis, the reader may consult the articles by Burt, Gaver, and Perlas,[3] Burt and Garman[2] Hartley and Wortham[6], and Martin.[7] For the sequential problem, it would be virtually impossible, even to construct an open-form expression for expected project completion time. This is due, in part, to the fact that the ordering of decisions is unknown à priori, and in part, to the fact that an optimization model must account for the dynamic behavior of the problem(3).

Because an optimal solution of this problem is unobtainable via analytic methods, several heuristic rules were developed for making the allocation decisions. The time required to complete the entire project, given that a particular heuristic is used for decision making, is a random variable. Estimates of the expected value of this project completion time were obtained by repetitively simulating "realizations" of the project. These mean value estimates (one for each heuristic studied) were then used to make comparative judgements of the effectiveness of the different heuristics in minimizing expected project completion time.

In the next section, four heuristics for making the resource allocation decisions are discussed. In actuality, only two distinct heuristics are presented, each of which was applied both statically and sequentially. The first heuristic is based upon the deterministic assumption that each activity will require an amount of time exactly equal to the expected value of the time to complete that activity. The second heuristic is an extension of the first which incorporates information reflecting the stochastic nature of this problem.

DECISON RULES FOR ALLOCATING RESOURCES

RULE I: Deterministic - Static

Suppose we make the assumption that the time required to complete any activity, given some resource level, is equal to its mean (or expected value). The time to complete activity i is then,

$$E\left[200\ (1-X_i)\ R_i\right] = 200(1-X_i)E\left[R_i\right] = 100(1-X_i)$$

The problem being studied is now deterministic. Letting T denote the time to complete the entire project, we may reformulate problem P1 as,

P2: Determine X_i, $i = 1,2,\ldots,15$ so as to

Minimize: T

Subject to: $\sum_{i \in \{j\}} 100(1-X_i) \leq T$ for all j

$$\sum_{i=1}^{15} X_i^2 \leq 2$$

$0 \leq X_i \leq 1$ for all i.

This is a non-linear programming problem and may be readily solved by one of the existing codes for these problems. Rule I is to apply this procedure to the static version of problem P2.

This reader should note that Rule I is truly a heuristic in that it does not yield an optimal solution. The assumption underlying P2 destroys the essential stochastic character of the problem. The value of T associated with the "optimal" solution of P2 will always be less than the actual expected value of project completion time. This is due to the fact that the maximum of the expected values of several random variables (i.e. path times) is always less than or equal to the expected value of the maximum of those random variables.

RULE II: Deterministic-Sequential

The second rule is to apply Rule I sequentially over the duration of the project at those points in time when the predecessors of any node are completed. The programming of the procedure to accomplish this is rather involved because the structure (fixed versus variables X_i's) of the resulting non-linear programming problems changes. The procedure runs something like this: Initially, problem P2 is solved with all X_i's variable; then the resource variables X_1, X_2, and X_3, egressing from node 0, are fixed (i.e. decisions are taken and resources are allocated); then R_1, R_2, and R_3 are used to determine whether the next most imminent decision will occur at node 1 or node 2 (4); Suppose the next decision will occur at node 2 (i.e. the actual time to do activity 2 is less than that of activity 1) -- then a new non-linear problem is solved with X_1, X_2, and X_3 fixed and with actual (rather than expected) activity times for activities, 1, 2, and 3; problem P2 is solved again and the resulting optimal values of X_6, X_7, and X_8 are fixed since those resources are now committed; then the next most imminent decision is determined, and so on. For each realization of a single project completion time, six non-linear programming problems are solved sequentially.(5) Rule II is simply the application of this procedure.

(3) Theoretically at least, it would be possible to formulate the sequential problem as a dynamic program with stochastic state variables by conditioning upon the different possible orderings of decisions (there are 22 such orderings for this network).

(4) The next most imminent decision can not occur at node 3 because it is "preceeded" by node 2.

(5) Although there are eight nodes in this network, only six programming problems, need be solved at each realization. This is because the optimal initial decisions (fixing of X_1, X_2, and X_3) are independent of the random numbers, and the optimal final decision (fixing of either X_{11}, X_{13}, or X_{15}) is simply to exhaust the remaining budget.

RULE III: Stochastic-Static

Rules III and IV are extensions of the two rules above, in which some "stochastic information" is used to modify the deterministic results. For instance, suppose Rule I has been applied to the problem so that the allocation decisions X_i, $i = 1,2,\ldots, 15$, are known. N simulation experiments can then be performed using these X_i's by taking random draws, R_1 (k), R_2 (k), ..., R_{15} (k); k = 1, 2, ..., N. (Note that this is the same procedure that is used to determine the estimate of project completion time achievable with Rule 1.) These simulations provide a method for estimating activity and path criticalities.[6] These criticalities provide information that may be used to modify the original set of X_i's in such a way that the expected project completion time is reduced (below that achieved with Rule I). For instance, suppose we consider a pair of activities, i and j, whose original resource allocations are X_i and X_j, and whose criticalities (estimated via simulation) are p_i and p_j, respectively. Suppose that $X_i < X_j$ and $p_i > p_j$. This means that activity i was allocated less resource than activity j, yet i is more likely to lie on the critical path than j. Ceteris paribus, expected project completion time could be reduced by taking some resource from activity j and giving it to activity i. Unfortunately the ceteris paribus condition does not hold because a modification in any of the X_i's tends to change all the other activity criticalities. However, as a heuristic the following simple procedure proved to be successful.

Step 1: Given a "current" set of X_i's, perform N simulations of project realizations so as to estimate the activity criticalities, p_i.

Step 2: Amongst all activity pairs (i,j) such that $X_i \leq X_j$ and $p_i \geq p_j$, determine the particular pair (i^*, j^*) which maximizes the expression $(X_j - X_i) + (p_i - p_j)$.

Step 3: Reduce X_j^* and increase X_i^*. The associa- cost reduction of allocation j^* must equal the cost increase of allocation i^*, so that the budget constraint remains an equality.

Step 4: If the current value of the expression $(X_j^* - X_i^*) + (p_i^* - p_j^*)$ falls below a predetermined level, then stop; otherwise, return to Step 1.

There are three important issues relating to the parameters used in applying Rule III. The first is the question of the "best" sample size, N, to be used in the simulations that estimate the criticalities. Too small an N would lead to statistically inaccurate criticality estimates; too large an N would require excessive computing time. This question becomes particularly important when Rule III is applied to the sequential problem. For the experimental results reported below, a sample size of N = 500 was found to be satisfactory. The second question involves the amount of resource to be shifted between activities i* and j* in Step 3 above. Too small an amount leads to an excessive number of iterations before the procedure is terminated. Too large an amount leads to radical changes in the criticalities, thereby "jumping over" good solutions.[7] Experimentally, it was found that incrementing the resource allocation of activity i* by one fourth the maximum value of the expression in Step 4 worked fairly well. The final question has to do with the termination level of the procedure in Step 4. A value of .08 appeared to yield a good tradeoff between excessive computing time and marginal solution improvements.

Rule III involved the use of the above heuristic procedure to derive a stochastically modified set of X_i's for the static problem.

RULE IV: Stochastic - Sequential

Rule IV is simply a sequential application of Rule III. As the predecessors of a given node are completed, all the unfixed X_i's are set via non-linear programming. These unfixed X_i's are then modified by performing a series of simulations as in Rule III. The simulations involve the use of actual times for activities associated with fixed X_i's and random variates for all other activity times. Then the X_i's egressing from the particular node under study are fixed and the procedure is repeated.

SIMULATION RESULTS AND ANALYSIS

Table I below presents the results of repetitively applying the four rules to obtain estimates of mean project completion time. The sample sizes used in determining the estimates vary because the amount of computing time required for a realization of the project completion time differs greatly between the static and sequential rules. Regardless of the rule being studied, each realization of project completion time required a set of 15 random numbers associated with the 15 activity times. For experiments on Rules I and III, the same 10,000 sets of random numbers were used so as to reduce sampling error. Likewise for Rules II and IV, the same 100 sets of random numbers were reused. Moreover for the latter two rules, the sample of 100 sets was in fact two "antithetic variate" samples, each of 50 sets.[8] The computing run time (exclusive of

(6) Definition: An activity's (path's) criticality is the probability that it will lie on (is) the longest path. Path criticalities may be determined via simulation experiments by counting up the number of times each path is critical and then dividing by the total number of experiments. Activity criticalities may be determined directly from the path criticalities.

(7) In preliminary experiments, the amount of resources to be shifted was quite large and this lead to cycling of the procedure. I.e., in one iteration of the procedure, resources would be given from i to j while in the following interaction, resources would be given from j to i.

compiling time) requirements for generating the specified number of realizations are shown for each rule; computing was done on an IBM 360, Model 91. The sample standard deviations of the mean estimates are given. Realizations in the different experiments are independent and hence the means should be roughly normally distributed. Various hypotheses about the significance of the differences in the mean estimates could be tested via standard statistical methods.

ity functions or activity cost-time tradeoffs) may be a considerably greater factor than this figure. Had the degree of improvements been 15 or 20%, then one would be justified in arguing the importance of using sophisticated allocation rules in project management. For the particular model and network studied in this paper, this was not the case.

	Rule 1: Determ.-Static	Rule II: Determ.-Seq.	Rule III: Stoch.-Static	Rule IV: Stoch.-Seq.
Sample Size	10,000	100	10,000	100
Mean Estimate of Project Completion Time	329.12	324.47	325.59	317.57
Standard Deviation of Mean Estimate	.4772	4.398	.4570	3.891
Computing Time in Seconds	2.66	337.16	3.81	613.15

Table I: Simulation Results

As was anticipated, the sequential rules led to shorter project sompletion times, on the average, than did their static counterparts. Similarly, each of the stochastic rules yielded an improvement over the results achieved with the corresponding deterministic rule.(9) These improvements are bought at a cost of computer running times which vary directly with the effectiveness of the different rules.

In terms of initial expectations, the most surprising aspect of the results is the relative closeness of the four mean estimates. The mean project completion time achievable with the most effective rule (Stochastic-Static) is only about 3 1/2% less than that of the worst rule (Deterministic-Static). Speeding up the completion date of a multi-million dollar project by 3 1/2% may be significant. However, the sensitivity of the results to inaccuracies in the data underlying the model (egs., assumptions about probabil-

(8) Antithetic variates is a Monte Carlo technique for reducing the variance of estimates obtained via sampling procedures. The standard deviations given in the Table for rules II and IV are unbiased values but they do not indicate accurately the "goodness" of the associated mean estimates; the actual standard deviation of those estimators are considerably smaller than the values shown. The reader may see Burt and Garman[1], and [2], or Burt, Gaver, and Perlas[3], for discussions of the effectiveness and application of Monte Carlo techniques in the analysis of stochastic networks.

In the opinion of this author, the closeness of the results for the different allocation rules is due, in part, to the following factors: the uniform probability functions, the quadratic costs, the deterministic assumption underlying the programming procedure, and the particular network configuration chosen for study. Each of these factors affects the degree of control that can be exercised over the processing of the project. If the probability functions had smaller variances (than those of the uniform), there would be a closer correspondence between the anticipated activity times achievable with a given resource allocation and the times that are actually realized. Under the uniform probability function assumption of this paper, suppose a manager allocates 1/4 unit of resource to a particular activity. Any time between 0 and 150 is equally likely to actually occur for that activity. Thus, there is a great deal of uncertainty as to how much will be accomplished (in terms of actual activity times) with a given amount of resource.

The second and third factors, those of quadratic costs and the deterministic assumption, has a very detrimental impact on the effectiveness of the non-linear programming algorithm to achieve

(9) The results in Table I may be contrasted with the mean completion times achievable with the following two naive rules: i) Allocate the budget equally amongst all activities (361.45) and ii) allocate the budget equally amongst activities 1, 2, 5, 6, 12, 14, and 15 (368.77) .

optimal allocations. The algorithm makes (sometimes very costly) resource allocations only to those activities lying on paths that are critical under the deterministic assumption. To illustrate the cause of resulting errors, consider the network of Figure II.

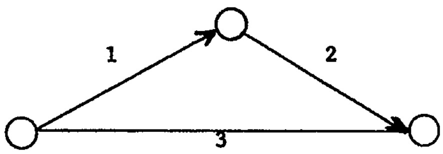

Figure II.

If a budget of slightly less than .5 dollars was to be allocated, activities 1 and 2 would each receive slightly less than .5 units of resource $(.5^2 + .5^2 = .5)$. Activity 3 would receive nothing since it is not critical under the deterministic assumption. The marginal cost of the last bit of resource put into activities 1 and 2 is extremely high compared with the marginal cost of zero for activity 3. Moreover, activity 3 will be critical only slightly less than 50% of the time. In an optimal solution to the problem of Figure II, activity 3 would receive some allocation of resources. The programming model of this paper leads to non-optimal solutions because of this assumption of deterministic activity times; furthermore, due to the quadratic nature of the cost functions, these non-optimal solutions may prove to be very costly (i.e., in terms of prematurely exhausting the budget).

A final factor that may have influenced the results of these experiments is that of network configuration. The basic concept is easy to illustrate with the aid of Figure III.

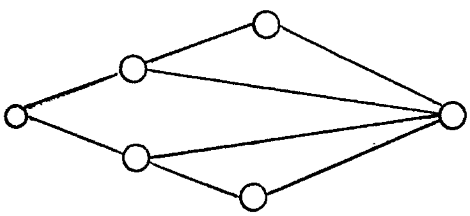

Figure III.

If the activities indicated in the figure were carried out from left-to-right, the manager would be able to sequentially "control" the project more effectively than if the activities were carried out from right-to-left. In the former case, resources would be committed initially to only two activities, while in the latter four commitments would have to be made initially. When processing is performed from left-to-right, the project manager is able to delay more allocation decisions, pending forthcoming information on realized activity times, thereby maintaining greater flexibility of control. The network studied in this paper has an equal number of arcs entering the source (Start) and sink (Finish) nodes. However, the structure of the entire network is such that a greater amount of resources must be committed initially (e.g., activities 1 and 2 utilize about 1/4 of the entire budget) when processing is done from left-to-right rather than in the reverse direction

Each of the factors mentioned above are subjects for research on the behavior of stochastic networks. In the final section below, some comments on directions for future study are given.

EXPERIMENTAL DESIGN ALTERNATIVES AND FUTURE RESEARCH

Any model for analyzing the stochastic, sequential, resource allocation problem must specify numerous assumptions about the structure underlying the system under study. The following partial list indicates some of the dimensions along which these assumptions may vary.

1. Network size and configuration.
2. Probability functions to represent the affects of resource allocations on activity times.
3. Choice of decision nodes; which of the activities are subject to the project manager's control?
4. Cost-time tradeoffs for each activity; are the quadradic cost assumptions of this paper realistic?
5. Objective functions; is the assumption of myoptically minimizing project completion time reasonable?
6. Resources; should we consider the cases of multiple resources, unconstrained resources, non-"exhaustible" resources, and non-continuous resource allocations?
7. Allocation decision rules; what heuristics might lead to good problem "solutions"?

The list could go on almost indefinitely. Even these few dimensions should give the reader a feel for the scope of this problem. For the purposes of this study, many simplifying assumptions were used such as those of using identical forms for activity costs and activity time probability functions. This was done primarily so that attention could be focused on the comparative effectiveness of decision rules. If other, more realistic data were available, the model and solution procedure could be applied with only minor modifications.

As mentioned earlier, one major area of future research has do do with identifying the relative "importance" of the different dimensions above with respect to their impact on the project manager's ability to achieve the specified objective. A second area of research concerns the realism of the assumptions dimensioned above and upon the data requirements underlying those assumptions. Are manager's concerns about uncertainty and sequential decision making as wisespread as was implied in the introductory section? Do managers think in probabalistic terms such as "expected completion time"? Even if it is very convenient for theoretical models, is data obtainable which could be used to construct the activity time - resource level probability functions? Greater attention must be given to field study research and behavioral information so that the management scientist can construct realistic models.

A third direction for research, which this author is pursuing, is the use of analytic methods to optimally solve the problem of this paper for relatively simplistic networks. For a project composed of a single serial path, dynamic programming gives such a solution. By using conditional random variables, dynamic programming may be applicable to more complex networks such as series-parallel-series configurations. Results of this analytic study may provide insights or new heuristics for treating large-scale problems.

This paper has presented a model for studying the problem of sequential resource allocation to projects in an uncertain environment. Due to the stochastic nature of the problem, analytic procedures could not be utilized to determine optimal solutions. Simulation provided a method for analyzing the comparative effectiveness of allocation rules, and results were presented for a particular set of problem specifications. The study was investigatory in nature and some directions for future work were indicated. Hopefully this paper will assist both theoretists and practioners interested in project management systems.

Acknowledgment: The author wishes to express his appreciation to Mr. Henry Garrett for his assistance in computer programming, and to Professor Glenn Graves for the use of his non-linear programming code.

REFERENCES

1. Burt, J.M., and Garman, M.B., "Monte Carlo Techniques for Stochastic PERT Network Analysis," Proceedings of the Fourth Conference on Applications of Simulation, New York, 1970.
2. Burt, J.M., and Garman, M.B., "Conditional Monte Carlo: A Simulation Technique for Stochastic Network Analysis," Management Science, Vol.18, No. 1, September, 1971.
3. Burt, J.M., Gaver, D.P., and Perlas, M., "Simple Stochastic Networks: Some Problems and Procedures," Naval Research Logistics Quarterly, Vol. 17, No. 4, December, 1970.
4. Davis, E.W., "Resource Allocation in Project Network Models - A Survey," The Journal of Industrial Engineering, Vol. 17, No. 4, 1966.
5. Elmaghraby, S.E., Some Network Models in Management Science, Springer-Verlag, New York, 1970.
6. Hartley, H.O., and Wortham, A.W., "A Statistical Theory for PERT Critical Path Analysis," Management Science, Vol. 12, June, 1966.
7. Martin, J.J., "Distribution of the Time Through a Directed Acyclic Network," Operations Research, Vol. 13, 1965.
8. Van Slyke, R.M., "Monte Carlo Methods and the PERT Problem," Operations Research, Vol. 11, 1963.
9. Vazsonyi, A., "L'Histoire de Grandeur et de la Décadence de la Méthode PERT," Guest Editorial in Management Science, Vol. 16, No. 8, April, 1970.
10. Wiest, J.D., and Levy, F.K., A Management Guide to PERT/CPM, Prentice-Hall, Englewood Cliffs, New Jersey, 1969.