Creative Construction Conference 2019, CCC 2019, 29 June - 2 July 2019, Budapest, Hungary

Economic Based Limited Resource Scheduling Algorithm

J. Mark Taylor, PhD, JD

Auburn University, 232 Gorrie Center, Auburn, AL 36849, USA

Abstract

Activity networks have been widely used in the planning, management, and control of building construction projects since their introduction in the 1950s. As the popularity of activity networking procedures has grown, so has the theory underlying the various procedures. Procedures are presently available that consider both technological and non-technological constraints and offer solutions solved optimally or heuristically. However, activity networks typical of building construction projects have not lent themselves to efficient optimal solutions because of their size, complexity, and diversity. Therefore, various heuristic scheduling procedures have been developed to solve this class of problem pragmatically if not optimally.

A disadvantage to the heuristic scheduling schemes available is that they follow a set of rigid heuristics and therefore may not be sensitive to a variety of differing network types or a variety of constraints. This study undertakes the development of a scheduling procedure that is heuristically based but that is sensitive to any variety of network types or constraints imposed by the scheduler. This is accomplished by combining all pertinent factors through a utility analysis and scheduling each activity based on the results of this analysis. The factors that are considered pertinent for the purpose of this study are the activity times based on technological constraints, network complexity, resource usage and availability, resource unit cost, and total activity cost. These factors are combined through the use of a cumulative utility model and prioritized to yield a scheduling sequence.

1.   Introduction

Activity networks typical of building construction projects consist of a series of activities describing the technological sequence of events necessary for the completion of the particular construction project. Since construction projects are typically non-cyclical operations, the activities contained in the network are non-repetitive in nature [1]. The activities may also be described as unique, as they describe a number of different work activities in a number of different geographical locations, and they consume a variety of scarce resources in various amounts. The networks may also be quite large, consisting of hundreds or thousands of independent activities. As a result of the diversity of the activities and of the magnitude of the activity network, optimal solutions have not proven effective. Heuristic procedures have been used in place of optimality in order to achieve adequate results [2, 3].

The focus of this research is to evaluate the feasibility of scheduling a construction project based on any number of considerations that could impact the efficiency of the schedule. The underlying premise with a scheduling procedure is based on an ordinal utility ranking [4, 5, 6] that results in scheduling conflicts that will be reduced to a minimum if the activities with the highest utility values are scheduled first. It is important to keep in mind that the highest utility values are indicative of activities that require the most restrictive resources, are the costliest, and/or are the most

restrictive in their placement within the network diagram. If many of the scheduling conflicts are eliminated from the outset, then the scheduling procedure should be more efficient than those used for similar types of problems.

In order to efficiently schedule the activities, each activity is assigned a utility value. The utility value can consider the aforementioned items or can be modified to consider any number of conditions deemed important by the scheduler [7, 8].

2. Scheduling Methodology

In order to schedule activities sequentially based on a concept other than time, it is necessary to use a scheduling procedure that treats time in a likewise subordinate fashion. That is, a scheduling procedure that is not based on the standard concept of beginning at day one and proceeding through the schedule a day at a time, scheduling activities as the resources are available until all activities are exhausted. The scheduling procedure must be flexible in that it allows the scheduling or time placement of activities in a random fashion yet does not violate any of the inherent technological or resource constraints of the activities.

In developing a computer program for such a procedure, efficiency is obtained by performing a minimum of checks where possible. Much of the scheduling procedure is devoted to performing checks for resource and/or technological constraints, but such checks are not required for the majority of the activities as the conflicts do not occur because of the basic premise of the scheduling procedure.

3. Scheduling Procedure

The scheduling procedure begins after the utility value for each activity has been determined. Beginning with the highest-ranking activity (the activity with the highest utility value), each activity is scheduled or assigned an appropriate time. As the scheduling does not proceed sequentially with time, it is first necessary to increase all activity times by a certain amount to facilitate backward scheduling as well as forward scheduling. For this study a value of 100 was added to all activity times, which proved to be sufficient for all activity networks tested.

Another preliminary consideration is the method of checking resource conflicts. As each activity is scheduled a resource profile is checked to see if the necessary resources are available during the appropriate time period. To facilitate this, a resource availability array is established with the available number of resources for each time period. Maximum limitations used in this study are 12 resources and 1000 days. The resource availability array established for a sample network was initialized with 35 units of Resource 1, 40 units of Resource 2, and a single unit of Resource 3. As activities are scheduled, the level of resources is reduced accordingly.

Activities are scheduled, as previously stated, by beginning with the highest-ranking activity and considering one activity at a time until all are scheduled. If resources are available in sufficient supply the project duration remains unchanged from the initial activity times as established without regard to resource requirements. If, however, the level of available resources is restrictive, the project duration will be extended. As the resource level is lowered, the project duration must be increased as more and more parallel activities are restricted to a serial placement. The scheduling of activities can be broken down into four different phases: (a) checking for resource availability, (b) checking network constraints, (c) resolving conflicts because of resource or network constraints, and (d) establishing activity times for each activity.

4. Resource Availability

The first consideration in the scheduling procedure is the availability of resources in sufficient supply to complete the activity as planned. The resources used by the activity are checked in order of the supply limitations imposed by the scheduler; that is, the resource that is most limited is first checked, the second most restricted resource checked next and so on. This is done as the most likely resource conflict will occur with the resource available in the most limited supply and if this is checked first any conflict is immediately determined and further resource checks are unnecessary.

Another consideration is given to the scheduled time placement of the activity. The activity is ideally placed between the early start time and the late finish time. If the activity is on the critical path as initially determined, there is no

option as long as the required resources are available. The activity should be scheduled on consecutive days beginning with the early start time. If, however, the activity is not on the critical path there is some discretion as far as the time placement is concerned.

Most heuristic scheduling procedures will always attempt to begin an activity at its early start time. The reasoning, of course, is that if any conflicts occur it will be advantageous to know as soon as possible and alternate activities can be checked for possible scheduling. This is not a consideration if the scheduling procedure is not sequentially based on time.

The ideal placement of an activity is as close to its late finish time as practical. By doing so expenditures, for the activity are delayed and project funds are conserved. If activity time statistics are available as in a PERT analysis, a possible time placement for an activity would be one standard deviation from the late finish time as this would delay the start of the activity to a point where it would still be completed by the late finish time 95% of the time. As this study used a standard network analysis where activity statistics are not considered, a mid-point preferred placement time was selected. That is, an attempt is made to schedule the activity midway between its early start time and its late finish time. As activity scheduling can proceed either forward or backward from this point, the preferred placement time eliminates any bias in either direction.

If resources are not available in a sufficient supply midway between the early start time and the late finish time, the remaining time between the early start and late finish is checked for resource availability. The activity is scheduled within this time frame as close to the mid-point as possible. This, of course, is attempted as network constraints will not be in conflict if the resources required by the activity are available between the early start time and the late finish time. If resources are not available within this time frame, additional checks are required to locate a possible scheduling time for the activity.

The check for available resources outside the early start and late finish time frame is a two-part sequence as scheduling can proceed in either direction. First, beginning at the early start time and advancing in a forward direction a tentative scheduling time is determined. Next, beginning at the late finish time and advancing in a backward direction a tentative scheduling time is determined. The scheduling time that is selected is the one that is closest to the original early start or late finish time.

This process is repeated for each resource required by an activity. As each resource is considered, it is checked against compatible times established for the previous resources. At the completion of this phase of the scheduling procedure each activity is assigned a preferred time. This time is the closest time to the original early start-late finish mid-point in which all required resources are available.

5. Network Constraints

After all resources required by an activity are satisfied, the activity time established is checked against the activity times for all other work items in the network diagram. Stated another way, is the early start time established for this activity greater than or equal to the late finish time for all preceding activities and is the late finish time of the activity less than or equal to the early start time of all succeeding activities? If the activity times established are between the original early start and late finish times there will be no conflicts with other network constraints; if this is the case, the times for all activities within the network are simply updated to account for the specific time placement of the activity rather than the time range established initially.

If the time placement of the activity falls outside the rage of early start-late finish times, a conflict will occur with other activities in the network. If the conflict occurs with an activity that has a lower utility value (e.g., an activity that has not yet been considered for resource availability and scheduling), the activity times for the lower ranking activities are updated; establishing new start and finish times for the activity.

By considering the activities for scheduling in sequence based on their utility value, most of the conflicts that occur are resolved by updating the activities that have yet to be scheduled. If, however, a conflict occurs with an activity

that has a higher utility value (e.g., an activity that has already been scheduled), this conflict must be resolved before further scheduling can take place.

6. Scheduling Conflicts

Scheduling conflicts that occur with activities previously scheduled are categorized into two groups. The first group consist of activities that are in conflict with a single previously scheduled activity and the second group consists of activities that demonstrate multiple conflicts.

In order to resolve the conflict that occurs between two activities, the scheduling sequence is reversed, between the two. This type of conflict would occur when Activity A was scheduled from day x to day y; and Activity C was scheduled from day x' to day y'. Activity B, with a lower utility value than A or C is to be scheduled as a successor to Activity A and a predecessor to Activity B and the duration of Activity B exceeds the difference between x' and y'. In order for activity A to be scheduled, one of the previously established constraints must be relaxed.

When this type of conflict occurs the resources that were previously assigned to Activity C are unassigned (i.e., returned to the resource pools) and Activity C is assigned a utility value just lower than the utility value of Activity B. That is, Activity A is scheduled, the activity times updated for B and C; Activity B is scheduled, the activity times updated for C; and finally, Activity C is scheduled without resource or network conflicts.

By reversing the scheduling sequence of only those activities in conflict, the remaining activities are unaffected and therefore the conflict is efficiently resolved. For example, if there are 100 activities in a network and a conflict occurs in the scheduling of the ninetieth activity, this activity and the conflicting activity are scheduled in reverse order without rescheduling other activities. If, however, there are multiple conflicts, this methodology would not likely result in efficiency as there would be a separate solution required for each conflict and these would have to be resolved simultaneously.

If multiple conflicts occur in the scheduling of activities, the activities in conflict are noted and scheduled before any of the other activities. If multiple conflicts occur repeatedly in the scheduling phase, the conflicting activities are successively added to the list of activities to receive primary consideration and the scheduling is restarted. As network complexity and resource restrictiveness increase, so do the incidences of multiple conflicts. The scheduling procedure as developed, resolved most multiple conflicts within a few passes of the scheduling sequence without adversely affecting the computational time required. The maximum number of multiple conflicts observed occurred in a highly resource constrained network. This network consisted of 97 activities and each required the use of eight resources. Forty-two multiple conflicts occurred before a solution was obtained and scheduling required more computational time of the same network with less restrictive resource requirements.

7. Activity Scheduling

Once all resource requirements are satisfied and all conflicts as a result of technological constraints resolved, the activity can be scheduled. By scheduling an activity, a specific time frame is assigned to the activity and the resources required for the activity are considered available only during this time. The initial early start, late start, early finish, and late finish times are therefore no longer valid. In place of these times is a start time and a finish time with the difference between the two equals to the duration of the activity.

As the times are assigned to the activity, the resources required are deducted from the available supply (e.g., the resource availability array) to prevent duplication of resource assignments. The activity does not change from this scheduled time unless a conflict is encountered as previously explained. The scheduling procedure is then repeated for the next activity in the utility ranking until all activities have been scheduled.

8. Results of a Scheduling Example

To illustrate how this process works a sample network of ten activities is used. The activities required the use of one or more of three resources limited in availability to 35, 40, and 1 unit per day respectively. Based on the resource requirements and limitations and the network constraints, the utility ranking is determined and used as a basis or sequence for scheduling.

The initial analysis of this network resulted in a project duration of 100 days but because of the resource restrictions imposed the duration was extended to 113 days. To schedule these activities only 10 iterations were required (i.e., no conflicts occurred with previously scheduled activities.

If this example is worked using standard resource allocation procedures based on early start times, as is typical of the most commonly used procedures, a similar schedule results. The major difference in the two schedules revolves around two activities. In the procedure developed for this study the activity with a high utility value is the second activity scheduled. The activity with a lower utility value is one of the last activities scheduled. This results in the scheduling of the first activity close to the originally determined activity times and the delay of lower ranked activity until adequate resources are available. This does not adversely affect the schedule as the first activity received a higher utility value be being on the critical path and the later activity has the maximum total float to work with.

Using a standard resource scheduling procedure, activities would be scheduled in a different order based, essentially, on time. This results in the delay of the project to a duration of 115 days which is slightly higher than the utility-based scheduling procedure.

9. Summary

In scheduling resource constrained networks, the final schedule is intimately dependent on the available level of resources. If all resources are available in unlimited supplies or more specifically in any degree required by a particular schedule, then the scheduling is effectively accomplished by an initial forward and backward pass. If the level of available resources is reduced to a point in which activities are forced to compete for the use of the resources, the scheduling of the activities becomes more critical if the original project duration is to be realized. That is, the activities must be assigned specific rather than general time placements in order to avoid delay causing resource conflicts. If the level of available resources is further reduced, the only option is to extend the project duration (assuming of course that the scheduling methodology is not to be changed). All resource constrained network scheduling procedures will accomplish this with varying degrees of accuracy.

There are basically two types of solutions to this type problem: optimal solutions and heuristic solutions. The optimal solutions have been shown to be either impossible or impractical for the majority of problems encountered with 'real world' construction scheduling. The heuristic solutions offer workable alternatives to the optimal solutions. The heuristic schemes deal with a comparative analysis of one scheme to another, that is, which is the best heuristic procedure - one based primarily on early start times, on late start times, on total float, on resource usage, etc. The answer of course is not general in nature and cannot generally be applied to all schedules. The important point is that each of these heuristic schemes is based on a non-varying set of heuristics. The procedures are not sympathetic to the specifics of the particular schedule.

By using a heuristic procedure that uses as a basis a criterion based on a combination of pertinent factors (i.e., utility value), an attempt is made to produce a scheduling procedure that is sympathetic to the peculiarities of the schedule. This is evident in that if a particular resource becomes more or less limited in supply, or if a particular resource becomes more or less expensive, or if a particular network diagram exhibits more or less complexity the sequence of scheduling these activities will change and, therefore, offer a different schedule. However, a difference in scheduling sequence may or may not offer a more favorable result. There is no heuristically determined scheduling sequence nor a basis for a scheduling sequence that results in the most efficient schedule for all network and resource considerations.

References

[1] Davis, E. (1966), Resource Allocation in Project Network Models: A Survey, Journal of Industrial Engineering, Vol. 17, No. 4, pp. 177-188. https://doi.org/10.1016/S0305-0548(97)00055-5
[2] Cooper, D. (1976), Heuristics for Scheduling Resource-Constrained Projects: An Experimental Investigation, Management Science, Vol. 22, No. 11, pp. 1186 - 1194. https://doi.org/10.1287/mnsc.22.11.1186
[3] Davis, E. and Patterson, J. (1975), A Comparison of Heuristic and Optimal Solutions in Resource-Constrained Project Scheduling, Management Science, Vol. 21, No. 8, pp. 944-955. https://doi.org/10.1287/mnsc.21.8.944
[4] Page, A. (1968), Utility Theory: A Book of Readings, John Wiley and Sons, Inc., New York, 1968.
[5] Seigel, S., et.al. (1964), Choice, Strategy, and Utility, McGraw-Hill Book Co., New York, 1964.
[6] Von Neumann, J. Morgenstern, O. (1953), Theory of Games and Economic Behavior, 3rd ed. Princeton University Press, Princeton, N.J.

[7] Weist, J. (1964), Some Properties of Schedules for Large Projects with Limited Resources, Operations Research, Vol. 12, No. 3, May-June, pp. 395-418. https://doi.org/10.1287/opre.12.3.395
[8] Weist, J. (1967), A Heuristic Model for Scheduling Large Projects with Limited Resources, Management Science, Vol. 13, No. 6, pp. 359-377. https://doi.org/10.1287/mnsc.13.6.B359

Note: References are limited to background material, as the focus of the paper is original work of the author and to his knowledge has not been published before.