Development of Software Tool for Scheduling Risk Analysis

Waleed M. Rashideh\textsuperscript{a*}, Ayham M. Alhazaimeh\textsuperscript{b}

\textsuperscript{a}Department of Information System, College of Computer and Information Sciences, Al Imam Mohammad Ibn Saud Islamic University, P.O. Box 5701, Al Riyadh, Kingdom of Saudi Arabia
\textsuperscript{b}Faculty of Computer Science and Information Technology, Software Engineering Department, University Malaya, Kuala Lumpur, Malaysia
\textsuperscript{a}Email: wmrashideh@imamu.edu.sa, \textsuperscript{b}Email: ayham.alhazaimeh@gmail.com

Abstract

There is a need in software industry for a tool to integrate scheduling risk. This tool will answer a multitude of problems in achieving a higher degree of scheduling accuracy. Successful project requires effective specialized tool for reducing risk. The objective of this study is to develop a project scheduling tool for the estimation of project risk. This study provides empirical evidence that scheduling tool are very much relevant, valid, and applicable in software industry. The output of this study is a tool which employs some of the main distribution methods such as Uniform, Beta, Triangular and Gaussian in order to decrease the risk of project delays. This tool was tested and the results were extracted, analyzed and discussed. The tool was developed using JAVA. It can be recommended that using a different type of probability distributions on project’s activities will give the project schedule more accurate estimations by using the developed tool early in the project development lifecycle.

Keywords: Scheduling; Risk; Analysis; Project; Approach; Distribution; Tool; Development.

1. Introduction

Project Time Management, also known as Project Scheduling, is the process of deciding the completion date for project. This activity is a part of the project planning. Project timing will set the project schedule. This step is one of the most challenging aspects project managers face due to the uncertainty in completion date. The fact is that most of projects slippage duration occurs early in project lifecycle. Current approaches and methods deployed to deal with this issue include methods such as, (CPM, PERT, and Monte Carlo simulation together with PERT).

* Corresponding author.
These methods remain partially incomplete and produce project schedules with different scales and results. Therefore, the obtained results from these methods may not enable the managers to succeed in achieving their project scheduling tasks. Although earlier studies have investigated this issue, however, it still remains a problematic area of project management. The study developed a software tool to help mitigate the risk associated with project scheduling. The tool can be used by work team to analyze risk associated with project scheduling. In addition, can be used to forecast the finish date of the project, thereby providing a decision tool for project managers in accepting or rejecting project task completion dates. The organization of this paper is as follows. Section 2 reviews some literature relevant to schedule risk analysis. The current scheduling risk approaches are presented in Section 3. Section 4 presents the used in scheduling risk analysis. A discussion is presented in Section 5. The last section of the paper describes the conclusions and several potential research directions.

2. Literature Review

Project management tools have been developed along with other powerful techniques to enhance the effectiveness of the project life cycle [1]. Project management techniques have also been considered as efficient approach to oversee and organize key systems development by many organizations. In short, project management techniques are powerful in managing projects especially software development projects.

Schedule constraint is an important dimension in project management. It may affect other constraints positively or negatively. In terms of negative impact on cost, it may increase the cost due to the delay. On the other hand, it may have positive impact on the degree of submitting the project deliverables. It is important to understand what schedule risk is all about. Schedule risk is the risk resulting from shortening or extension of project schedule [2].

Another factor that requires clear definition is time estimation. Software development time estimation is the process of estimating the most realistic use of time required for developing software [3]. In other words, it is used to determine the view of future; as a consequence, it is rare to have accurate estimation of activity duration [4]. The following section investigates and elaborates on different scheduling approaches.

The world facing rapid changes in either technology or businesses. These changes affect the organization on how they perform their businesses. Risk assessment is one of the core activities in industry [5]. As time management is one of the most important constraint in project management that has a direct impact on the cost of the project. This area facing highly risk in missing deadline that will have several negative impact on project and organization. Several scholars define the schedule risk as the delay of closing the project in time[6,7,8]. None of these definition shed the light on the impact or consequences of delay [9]. This shows the importance of risk scheduling technique. To have a clear picture of current approaches of scheduling techniques, the following section highlight these approaches.

3. Current Scheduling Approaches

Scheduling techniques are important for some projects depend on its type and nature. Some of the most utilized scheduling techniques are the Critical Path Method (CPM) and Project Evaluation and Review Technique
(PERT). These tools have certain limitations; and this study will enhance these techniques and utilize them in an efficient way. This research will then compare this technique with other tools implemented using linear scheduling, genetic algorithms, and other simulation techniques.

Fischer and Aalami [10] argued that the current approaches for scheduling need to translate the design information into tasks; and expressed the need for taking the advantage of the increase in "electronic and object-based descriptions of designs, schedules, estimates, and integration mechanisms that translate design descriptions into schedule view of projects" [10]. Computer-interpretable models are presented as one of such mechanisms. These models should be used to help in generating automated schedules.

Scheduling is important and necessary process in all phases of a project. CPM and PERT are the most popular methods used for scheduling compared to other approaches. This Research is intended to improve the computing effectiveness of these approaches as required in the dynamic requirements of the project development. In fact, this paper does not only discuss these two techniques but also evaluates other available techniques in the market.

Risk analysis tool is developed to answer some very crucial questions in the project with certainty and the potential impact of deviation. Schedule risk means that there is a chance of the project to exceed its estimated plan. To shed the light on the schedule risk analysis, a graph of schedule risk for project is shown in Figure 1. With a project has 40 months duration, the analysis gives the chance to complete this project within the time of around 80%; that is, a chance to finish the project in 34 months, with 50% (42%) certainty that the project will be completed in 68 months.

Ben-Haim and Laufer [11] designed a model for enhancing the project schedule reliability which faces uncertainties of its task duration. The research presents that the technique for implementing the concept needs minimum information that is easy to use and support in the construction of schedules at an acceptable level of reliability. Wang and Demsetz [12] have simplified the NETCOR (Model for Evaluating Networks under

![Figure 1: Nominal Schedule Risk](image-url)
Correlated Uncertainty), used for assessing project schedule networks when activity durations are related. The research presents the practical implementation of NETCOR on a project. PERT and other simulation analyses techniques that do not consider correlation were also assessed. Based on comparing the results, there is a significant of taking into consideration the correlation in scheduling process. The following sections explain in details the most popular tools for scheduling in project management field.

The increase in complexity of modern projects emphasizes the needs for efficient tools to meet the scheduling challenges faced in project management areas [13]. These challenges represent constraints on project success. The high rate of project failure dictates the need for investigating some of the issues with regard to project failure. Project expenses are observed to be increasing tremendously. Huge amount of this expenditure goes to technology, goods, services, and staff. Lately, IT projects have been observed to have terrible track record. Therefore, meeting the project scope, time and cost are considered to be the main constraints to project success. Several studies found that 44% of software projects met their time schedule within the budget, and 24% of all software projects were terminated before their completion time schedule [14].

The project failure is attributable to scheduling problems, resulting to a risky situation in software development [15]. In this case, risk management practices have many benefits. In addition to helping in anticipating problems, risk management improves the timely delivery of software project and reduces schedule slips [16]. Effective uses of these tools have helped in improving the success rate of software development projects with regards to scheduling problems. The tool developed in this project will enhance the scheduling risk analysis in estimating project risk. The following sections present the techniques adapted in this study.

3.1 Gantt Chart Technique

It’s a quantitative tool used in project management for scheduling analysis developed by Henry Gantt in 1917. It employs a graphical representation for the project phases [17]. This graphical diagram will give the project tasks the easy way to follow up and see where each task should be, and to show it’s the current project status.

As any other technique, The Gantt chart has some serious disadvantages when it is used for managing projects. It does not show the relations between project’s tasks. It is sometimes complex or has big projects. A lot of tasks will have pre-requirements in situations that need other tasks to be fully or partially finished before they can be started. To show these relationships, in case of a large number of tasks is not sufficient to be represented in a Gantt chart diagram [18].

3.2 PERT Technique

PERT technique contains a detailed diagram for all the project’s planned tasks [19]. These tasks’ list, with all of their precedence tasks, is shown in an organized network. Moreover, project managers would assign an estimation of probability distribution for task duration which can be taken from expert. In order to apply the PERT technique, the experts were asked to predict three estimations for tasks to be done. This is known as a three point of estimation i.e. the pessimistic which means the maximum duration for the task to complete, optimistic means the minimum duration, and most likely which will be a value between the previous two
durations [19]. Along with some other mathematical techniques, it is possible to compute and expect a probability distribution for the time to complete the whole project.

Experts used to see PERT as a very useful tool to estimate the life cycle of the project. Based on the success of PERT technique, there was an adaptation of PERT such as PERT/cost that used to address the project cost issues as well. While PERT was widely much-trusted by the business in the 1960s, later studies on the other hand mentioned some doubts regarding the success of applying PERT in the project management process [19].

PERT is developed using time as key factor, also its implementation is proved with saving two years from the actual expectation of five years required to finish the project. An important dimension of PERT is the use of three completion time estimates for each task, and the implementation of statistical probability theory to estimate the likely probability of finishing the project within a given time. The success of this implementation encourages the U.S. Department of Defense to determine that all contracts must use PERT.

3.3 CPM Technique

DuPont developed a similar planning and management technique during the same period that PERT was implemented. The Critical Path Method known as (CPM) that use a graphical depiction such as PERT, however, at the early stages, CPM task durations are not estimated based on probability distributions. CPM used to generate the base-line schedule [20]. The nature of the network representation allowed for easier computation as it also makes the computation of the critical path very easy, this means the tasks that drove the final project duration [17]. Many Improvements to this technique allowed the exploration of alternative resource allocations for reducing this time and cost of the project.

Scheduling process has many techniques employed to estimate the duration in the project life cycle, such as (critical path method) CPM. This method is used to calculate the shortest project completion duration through a network. The delay on the critical path will affect the completion date of the project. On the other hand, CPM presents these paths that are not critical can be extended if they have sufficient flexibility, without delaying the project [17].

Unlike other techniques, the (CPM) is well accepted among project managers. It is important for managing the daily activities and developing the logic of the project work. The CPM schedule is only accurate if every activity starts as planned, and this is considered as one of its drawbacks. Of course, project managers recognize that projects do not go completely according to pre-plan. CPM considers the starting point of project scheduling process [20]. CPM is single point prediction, and therefore gives a wrong idea that the future can be estimated correctly [17].

3.4 Monte Carlo Approach

One does not need deep knowledge of probability theory and statistics in order to carry out risk analysis. All that is needed is to understand that the probability value can affect decisions and that you can calculate the variability of a quantity as a simple range. Certainly, it is not always appropriate to perform a schedule risk
analysis; however, if there are different numerical uncertain values, using risk analysis tools is strongly recommended [21].

Highly complicated approach such as Monte Carlo simulation can be applied to projects when information is not exactly defined in advance.

Monte Carlo simulation is designed for accurate implementation and, as such, generally requires great accuracy in data inputs and well-defined probability density functions for the various variables [22]. Comparing cost and time data on engineering projects is lack of accuracy.

They are level of acceptance of variances of 5 to 10% when taking decisions related to risk quantitative values.

The key factors are probable at their unfavorable values and the contact is estimated. The outcomes are inescapably atrocious.

They are hypothetically in the worst scenario. The mathematical accessibility is at best ineffective and at worst, they are ambiguous.

Even in a highly risky project no variable will go to its unfavorable acute value. Some variables will state little discrepancy from the plan and other variables may convert into the favorable direction. Thus, the real contact compared to the worst case would be less [22].

Monte Carlo Simulation can determine the real actual situation and the probability of project’s success.

If applied a risk analysis and told management that they have an 80% probability of success and have a 45% exposure, what does it mean? It means the bottom line decision factor can erode by as much as 45% of the targeted value”.

The reason behind that is they want to look at a ranked list of risks and opportunities. The alternatives proposed then to be examined in the Monte Carlo environment to get an optimal result.

Monte Carlo can be used to evaluate a current plan and it can also be an excellent technique for expanding a better one [22].

In Monte Carlo analysis, the process of iteration (a single simulation run) can be performed in the following steps:

- For each random factor, select a sample based on the estimation distribution function and count the value;
- Run a simulation by using a random input samples and calculate the corresponding simulation result;
- Re-do the previous steps \( n \) times; and
- The \( n \) values from the iterations will determine the output probability distributions.
3.5 Probability Distributions

In a world of risk, outcomes are described according to their probability of occurring. A probability distribution shows the possible range of outcomes. Here is an example:

![Probability distribution](image)

**Figure 2**: Probability distribution

As shown in figure 2, probability distributions are chosen based on their ability to model the real world. In fact, a limited selection of distributions is useful for the vast majority of risk work. Even if you do not have the need to choose distributions yourself, knowing what is available will help you better visualize the nature of risk [23].

3.5.1 Triangular distribution

This is one of the Mathematical Distribution tools as shown in Figure 3. This method uses minimum, maximum, and the most expected values. It has been investigated by Johnson [24]. The triangular probability distribution is well applicable in models which lack of data. The bounding cases where $Minimum\_duration=a$ and $Maximum\_Duration=b$ are named right triangular and left triangular distribution respectively.

![Triangle distribution](image)

**Figure 3**: Triangle distribution

3.5.2 Normal distribution

Normal distribution also known as Gaussian is used to draw events that have equal probability of occurring above or below an average. Results are focused near this average, the farther away, the less likely is an
outcomes. Figure 4 shows the shape of the normal distribution.

![Figure 4: Normal distribution](image)

### 3.5.3 Uniform Distribution

This type of distribution has no peaks and because of this, all outcomes between the highest and lowest possible values are equally likely [25]. This is as shown in the figure 5.

![Figure 5: Uniform distribution](image)

### 3.5.4 Beta distribution

Beta distribution resembles the normal distribution.

Its tails may be specified more readily than the normal and they may either be symmetric or asymmetric.

Many real world phenomena are biased towards the low or high end of possible outcomes; the beta distribution can capture this [26]. Figure 6 shows the shape for this distribution.

![Figure 6: beta distribution](image)
All the mentioned mathematical Distribution approaches can be used in the scheduling estimation. Research [27]. Therefore, experiment is a good method for this purpose. Table 1 summarizes some of these mathematical tools.

**Table 1: Different Mathematical Distribution Approaches**

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Shape</th>
<th>Typical Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Triangle</td>
<td></td>
<td>• a model when the minimum, most likely and maximum are declared without knowing the data of distribution shape</td>
</tr>
<tr>
<td>Normal (or Gaussian)</td>
<td></td>
<td>• Normal distribution that’s used in different software project estimation or modeling techniques; can be used to estimate the durations.</td>
</tr>
<tr>
<td>Uniform</td>
<td></td>
<td>• Uniform employed when there is missing of distribution data</td>
</tr>
<tr>
<td>Beta</td>
<td></td>
<td>• Its tails may be specified more readily than the normal and they may either be symmetric or asymmetric</td>
</tr>
</tbody>
</table>

This section will present a table of summary for the existing scheduling approaches and tools that are addressed in the literature review.

**Table 2: summary for the existing scheduling approaches and tools**

<table>
<thead>
<tr>
<th>Scheduling approach/tool</th>
<th>Specifications of the scheduling approach/tool</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gantt chart</td>
<td>Represents a graphical show for the project progress phases, each phase will vertically represent the duration of activities by a horizontal line together with a time scale, while current time is represented using a vertical line moving from the left side to the right side [17]. Predicts a three estimation for tasks to be done; this is known as a three point of estimation the pessimistic which means the maximum duration for the task to complete, optimistic means the minimum duration, and most likely which will be a value between the previous two durations [19].</td>
</tr>
<tr>
<td>PERT technique</td>
<td>CPM is used to generate the base-line schedule [20]. The essence of the network representation supports the easiest computation; it also makes the computation of the critical path very easy, this means the tasks that drove the final project duration [17]. A random variable will be taken as a sample based on the probability distribution function and count the value. It depends on running a simulation using the random input samples and calculation of the corresponding simulation results. Repeat the previous steps n times. The n values from the iterations will determine the output probability distributions [22].</td>
</tr>
<tr>
<td>CPM</td>
<td>Probability distributions are chosen based on their ability to model the real world. In fact, a limited selection of distributions is useful for the vast majority of risk work.</td>
</tr>
</tbody>
</table>

4. Methodology

4.1 The Software Tool
The tool proposed for the purpose of this study, named the Scheduling Risk Analysis Tool (STSRA), is designed to meet end user’s needs, in particular project managers for reducing project schedule risks. STSRA is stand-alone software designed for scheduling analysis for project development process. It will not totally take out the schedule risks but it could help the project managers to predict early the risks in order to apply a certain plans to avoid the delay or the project failure.

4.2 Key functions of STSRA

The main function of the tool is to assist project managers in evaluate the project duration in early stage of the project life cycle taking into account the risk around the project schedule. This tool can be used by the group of workers who are using the tool for analyzing project risk schedule in the project management to predict when the project will finish and to decide whether to take it or not. This could be a very important issue in the organizations and the business area they are working in. Mainly, there are two parts in the proposed tool namely; simulation and distribution type. The first part allows the user to decide the number of iterations that simulation will apply on a specific project, and the second part will let the user choose the best statistical distribution to be applied on an individual activity of the project. After that and during the second phase of this study, a simulation run using data will be conducted and the results will be compared and analyzed to verify the accuracy of the software. Project success depends on risk analysis tools. It will make project managers to be more reliable on developing a new project for the organizations they are working in. Also, it will enhance the risk analysis skills in the organization. With the software tool for scheduling risk analysis, software development will catch the key specifications in the project duration. This software can be used for the following functions:

- To assist project managers in evaluating the project duration in early stage of the project life cycle taking into account the risk around the project schedule.
- The tool can be used by the group of workers who are using the tool for analyzing project risk schedule in the project management to predict about when the project will finish and to decide either to take it or not.

The most important key for organizations’ success is lessons learned from old mistakes and to avoid them for new projects. The proposed tool has a main function related to the issue of identifying those risks related to the project schedule early in the project life cycle so that people will be aware of this in the future. It will help in reviewing projects that were earlier implemented so that they conduct or find out the new project possible directions of success. The scheduling risk analysis tool helps in predicting project duration in days based on the information entered by project manager who will be taking into account the issues related to experience and historical information from previous project. For project manager who wants to lead a successful project, the proposed tool will give him a hand to be more accurate about the completion date of the project and then to avoid project failure. It is common that projects exceed planned duration most of the time, and this could cost the organization a lot of money. By using the tool, the image will be clearer for the stakeholders of the project to be more accurate about the finishing date. So, all the team will manage the work to deliver projects on time.

The scheduling risk analysis tool will allow the user to create new project, save and edit projects easily. The
users for scheduling risk analysis will not face any problem because it has a lot of specifications that make this tool very usable even for novice users. The tool is developed using GUI principles so it meets users’ satisfaction and makes risk analysis a more interesting process.

4.3 Implementation of Software

The implementation approach employed for this research is object oriented methodology. This approach will help simplify the re-use of software components. Several criteria are employed to develop the system:

- Schedule risks
- Reliability to estimate project duration correctly
- Early warning for project team about the risk
- Create business case for new software project
- Employ historical data to be used for future work
- User friendly criteria

4.4 Software Design

The development of this tool depends on utilizing the four models namely; uniform, Triangular, Beta, and Gaussian distribution. The developed system has the following modules:

- Project module is used to save project data permanently.
- Activities module which provides the user with empty activity list to be filled by the user
- Precedence module that allows users to fill the precedence activity based on the project implementation scenario.
- Distribution module is providing the system with essential input in selecting the distribution of each activity.

The general algorithmic schema for the tool can be addressed as following: indentify activity list, dependencies, select probability distribution, specify parameters, and number of iteration for simulation. Figure 7 represent the activity array function.

```java
public void addActivity() {
    activitiesArray.add(new ArrayList<Object>());
    activitiesNumber++;
}
```

**Figure 7: Activities Array Function**

On the other hand, topological sort algorithm is employed for project path. It computes an ordering on a graph. It is used to find the longest path in the project activity list. Figure 8 shows the topological sorting algorithm.
The topological sort algorithm is one of the most useful algorithms in this area. It computes an ordering on a graph such that if path ABC is earlier than path DEC. The simplest topological sort algorithm is to repeatedly be applied on the project paths to find the longest path in the project activities list [28]. The topological sorting is the best in scheduling network the Algorithm to compute the Earliest Completion Time the following algorithm will do so:

First the project activities should be sorted topologically as:

- For each event j: initialize $EOT(j) = 0$
- For each event $i$ in topological order
- For each event $j$ adjacent from $i$:

$$ECT(i, j) = EOT(i) + \text{duration} (i, j)$$

The algorithm will contain the following data structures:

- A two dimensional array that gives the duration of the activities.
- An array that contains the activities sorted in topological order
- A two dimensional array that will give the Early Completion Time of each activity. It has to be initialized with 0.

After implementing the desired code, user can assign dependencies to activities in the appropriate field. If user wishes to add more precedence, a semicolon should be used to separate them.

The mathematical distribution type for the activity will be used to generate the expected random duration. The parameters of each distribution differ from each other. For example, normal and uniform distributions have two durations (Maximum, Minimum) values. While Beta has (Beta, Alpha), and Triangular has three values (Minimum, Most likely, Maximum). The input will determine the random expected duration of the activity. Simulation is used to achieve the objective of the tool. For each time in the simulation the activity will have a different random number generated based on the distribution type. After $n$ times the simulation will give the maximum expected duration for the project by calculating the mean value for duration which can be found by dividing the total durations over number of simulations. Figure 9 represents extracted code showing a simple

![Figure 8: Topological Sorting](image)
call to a function for computing total duration of activities.

```java
public void computeMaxTotalDuration() {
    for(int i=0; i<repetitions; i++) {
        computeDurations();
        totalDurations[i] = computePERT_CPH();
    }
}
```

**Figure 9: Compute Total Duration Function**

**Figure 10: STSRA Main Screen**

**Figure 11: Critical Path Node**

The critical path node in STSRA is used to show the user how many times an activity has been picked in the
simulation. It is useful to know the activities that participated most in the critical path so that need to pay more attention to them mainly because they have the highest risk in delaying the project. Figure 11 shows critical path node.

5. Discussion

The system is developed to employ different kinds of probability distribution approaches. The parameters represent time needed to complete the activity. The system employs three time durations plus total number of iteration. Assuming that project have 10 activities of software project, although does not matter even 100 activities. Each activity should take up to three different expected finishing times to be added to the experiment in order to run the simulation of the software. Figure 12 shows the results of the experiment for 10 activities.

![Figure 12: System Input Value](image)

Triangular and Uniform distributions method are recommended to follow by project managers because they offer a challengeable duration time. On the other hand, Gaussian distribution type may have the longest duration time, however, this is with less risk possibility of losing the deadline of the project.

<table>
<thead>
<tr>
<th>Distribution Type</th>
<th>Scheduling method</th>
<th>100 iteration</th>
<th>1000 iteration</th>
<th>10000 iteration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Max Duration in days</td>
<td>Standard Deviation in days</td>
<td>Max Duration in days</td>
</tr>
<tr>
<td>Uniform</td>
<td></td>
<td>28.871</td>
<td>1.81</td>
<td>30.55</td>
</tr>
<tr>
<td>Triangular</td>
<td></td>
<td>28.582</td>
<td>1.416</td>
<td>30.032</td>
</tr>
<tr>
<td>Beta</td>
<td></td>
<td>33.368</td>
<td>0.253</td>
<td>35.457</td>
</tr>
<tr>
<td>Gaussian(Normal)</td>
<td></td>
<td>28.8</td>
<td>1.734</td>
<td>30.3</td>
</tr>
</tbody>
</table>
6. Conclusion

Project scheduling involves the determination of project start and finish dates. Project schedule illustrates the project activities and an expected duration for each activity in order to calculate the project completion date. The current approaches and methods in use to deal with the project scheduling and evaluating completion risks including Critical Path Method (CPM), Program Evaluation and Review Technique (PERT) and Monte Carlo Analysis and Simulation, to name a few, produce different results and insights toward a project schedule. This may not therefore enable the managers to succeed in achieving their project scheduling tasks. The results will also require further analysis and evaluation to avoid and eliminate scheduling pitfalls and risks associated with the project activities. The researcher developed a project scheduling tool named STSRA for the estimation of project risk tool to mitigate this problem.

The developed STSRA will contribute and provide an input to the project managers when setting up scheduling for a project. This contribution is crucial for the success of the project management activity. The appropriate view of the results given based on using different scheduling approaches of different iterations will help managers to estimate the risks associated with their projects during the initial phase of the project.

References


