Buffer sizing approach with dependence assumption between activities in critical chain scheduling

Li Bie\textsuperscript{1}, Nanfang Cui\textsuperscript{2}, Xiaoming Zhang\textsuperscript{3}

1. Office 326, School of Management, HuaZhong University of Science and Technology, Wuhan, 430074, China. E-mail: billybrown1983@yahoo.cn. Tel.: (0086)15807157883.

2. School of Management, HuaZhong University of Science and Technology, Wuhan, 430074, China. E-mail: nfcui@mail.hust.edu.cn.

3. School of Management, HuaZhong University of Science and Technology, Wuhan, 430074, China. E-mail: cani_chang@163.com.

POMS 22nd Annual Conference

Reno, Nevada, U.S.A.

April 29 to May 2, 2011
Abstract: The Critical Chain / Buffer Management (CC/BM) methodology, proposed by Goldratt (1997), introduced the concept of buffers to protect the critical chain against the disruptions. The buffer sizes reflect the uncertainty in project duration estimation and affect the project scheduling performance. However, the most present buffer sizing approaches make the assumption that project activity durations are independent. In this paper, a method for determining buffer sizes with dependence assumption between activities is introduced. Specifically, the number of activities affected by the risk factor and the level of dependence parameter are integrated into the formulation. The method is compared to the root square error method (RSEM) by using an example in the Patterson data set. The results indicate that the RSEM may underestimate the buffer size when both of the above factor value is at a high level, while the suggested method can provide better protection in such circumstances.

Key words: Project scheduling; Critical chain; Buffer sizing method; Statistical dependence

1. Introduction

Goldratt’s Theory of Constraints (TOC) and its direct application to project management, known as the Critical Chain / Buffer Management (CC/BM), has been a popular and effective approach to project management. Since 1997 the publication of Goldratt’s novel Critical Chain (Goldratt, 1997), it received considerable attentions among practitioners, and several books and articles have already appeared on the subject (Newbold, 1998; Leach, 2005; Woeppe, 2006; Herroelen and Leus, 2001; Tukel et al., 2006; Rabbani et al., 2007; Blackstone et al., 2009).

In CC/BM, the critical chain is defined as the longest path which not only considers the precedence, but also takes the resource dependencies into account. The
length of critical chain decides the project due date. Furthermore, the concept of buffers is introduced to protect the critical chain against the disruptions that may occur during project execution. Buffers are calculated to reflect the uncertainty in the estimates of duration of tasks. In the single project environment, there are three types of buffers: feeding, project and resource. Feeding buffers (FB) are placed on all paths that feed the critical chain, in order to project the critical chain from potential delays by the feeding paths. A project buffer (PB) is positioned at the end of the critical chain to protect project due date from variation in the critical chain tasks. Resource buffers (RB) is different from the project buffer and feeding buffers in that it doesn’t occupy time in the project network. It is placed whenever a resource has a job on the critical chain in the form of the warning tool (Leach, 2005, Tukel et al., 2006).

The most common buffer sizing approaches are the cut and paste method (C&PM) and the root square error method (RSEM). Tukel et al. (2006) integrated project characteristics into the buffer sizing approaches, and introduced the adaptive procedure with density (APD) and the adaptive procedure with resource tightness (APRT). Long and Ohsato (2008) determined the project buffer size by computations with fuzzy numbers.

The two strengths of the C&PM are that it is simple, and it usually provides a large-enough buffer. However, the size of the buffer increases linearly with the length of the feeding chain. The method creates relatively long and impracticable duration buffer in long projects. It doesn’t also allow the project engineers to explicitly account for known variation in the tasks. Compared to the C&PM, the primary advantages of
the RSEM are that it allows to use known task variation and it won’t generate very large or very small buffer sizes based on the length of the feeding chain. But it may lead to underestimate buffer sizes for long chains. The second disadvantage of the RSEM is that it assumes that project activities are independent without considering interruptions imposed by external factors (Herroelen and Leus, 2001; Leach, 2005). The APD and the APRT generate smaller buffer sizes than the C&PM and the RSEM, and also consume higher percentage of buffer sizes. They are better choices when uncertainty of project is low and the project manager wants shorter project completion time (Tukel et al., 2006). The method of Long and Ohsato (2008) is a more appropriate alternative when activity durations are estimated by the subjective judgment of the experts due to the lack of historical data.

The above buffer sizing methods make the assumption that project activities are independent. However, besides precedence relations, resource requirements and resource availabilities, the interaction between activity durations has a major influence on the duration of a project (Herroelen and Leus, 2001). The interaction between activities may result from some risk factor, such as weather, labor skills, equipment, and management quality et al.. These risk-related activities may vary together so that the uncertainty of project duration increases. This impact will become more obvious in the more complex project environment. So these approaches can lead to the underestimation of the buffer size.

Some previous papers started to focus on the dependence between activities in the buffer sizing approaches. Leach (2003) present that the activity independence
assumption is not valid in practice, and the invalidation is resulted from the positive
bias that may systematically increase schedule performance relative to the plan, so he
suggest that the project buffer is calculated as the sum of the variation buffer and the
bias buffer. Based on the suggestion of buffer sizing method provided by Leach
(2003), Trietsch (2005) consider systemic bias of the project as a statistical
dependence issue between activities, and develop a mathematic model for computing
the variance elements and bias correction. However, his activities dependence model
is oversimple. For example, the factors that cause the statistical dependence between
activities aren’t quantized, the number of activities that are influenced by the same
factor in a project aren’t specialized, so that the buffer sizes resulted by his model
aren’t precise.

In our study, we analyze the effect of level of dependence parameter and number
of risk-related activities on project duration performance, and introduce a method for
determining buffer sizes with dependence assumption between activities by means of
integrating the two variates to the formulation.

2. Buffer sizing approach

2.1 The cut and paste method

The buffer sizing method is proposed by Goldratt in his novel Critical Chain
(Goldratt, 1997). He suggests to use 50% of the activities duration as the safety time
regardless of their uncertainty level, then sums up the safety time along the path and
sizes the buffer as half of the total. Tukel et. al. (2006) refer to this as the cut and paste
method (C&PM). Leach (2003) refers to this as the ‘50% of the Chain’ method.

2.2 The root square error method

The buffer sizing method is proposed by Newbold (1998). In this method, the uncertainty of each task duration is calculated as the difference of its safe estimate and its average (50%) estimate. The method sizes the buffer as the square root of the sum of the square of the differences for the tasks along the chain. Newbold (1998) also suggests that the standard deviation of each the task duration is one half of the difference, then the buffer size is twice as much as the standard deviation of the chain. Tukel et al. (2006) refer to the method as the root square error method (RSEM).

Let $U_i$ be the duration uncertainty of activity $i$, $n$ be the number of activities in the chain, $S_i$ be the safe estimate of activity $i$, $d_i$ be the average (50%) estimate of activity $i$. Then the buffer size is calculated as:

$$BufferSize = \left( \sum_{i=1}^{n} U_i \right)^{1/2}$$

2.3 Adaptive Procedure with Resource Tightness

Tukel et al. (2006) propose two buffer sizing methods. The first method integrates project resource tightness into the formulation. The authors argue that the task is more likely to delay if the total resource usage is close to the total resource availability. So resource tightness (RT) is calculated as the maximum value of each resource utilization factor, which is defined as the ratio of total resource usage to the total resource availability for each resource. The buffer size is calculated as:

$$BufferSize = (1 + RT)^* \left( \sum_i^{VAR_i} \right)^{1/2}$$
\((\sum_{i} VAR_i)^{1/2}\) is the square root of the sum of the variations of the tasks making up the chain, and it denotes the standard deviation of the chain.

2.4 Adaptive Procedure with Density

The second method takes project network density into account. The authors suggest that the task is more likely to delay as the number of task precedence relationships increases. Network density (ND) is reflected as a ratio of total number of precedence relationships to the total number of tasks. The buffer size is calculated as:

\[
\text{BufferSize} = (1 + ND) \ast (\sum_{i} VAR_i)^{1/2}
\]

\((\sum_{i} VAR_i)^{1/2}\) is the standard deviation of the chain.

2.5 The root square error method with fuzzy numbers

The buffer sizing method is proposed by Long and Ohsato (2008), and the buffer size is determined as the square root of the sum of the squares of the safety time estimated by fuzzy numbers. The safety time of each activities is calculated as the difference between the suitable deterministic duration and the high agreement duration in fuzzy number model of TrFN \((a,b,c,d)\). The high agreement duration is measure by so-called the agreement index (AI) which is defined as the percent of the fuzzy event A inside the boundaries of the fuzzy event B.

Let \(s_t\) be the safety time of activity \(i\), \(t_i^h\) be the high agreement duration, \(t_i^d\) be the suitable deterministic duration, \(P\) be the number of critical chain in the initial deterministic schedule, then the project buffer size is calculated as:

\[
st_i = t_i^h - t_i^d
\]

\[
\text{BufferSize} = max_{p=1..P} (\sum_{i \in P} st_i)^{1/2}
\]
3. Effect of activity dependence on project duration mean and standard deviation

As pointed out earlier, most existing buffer sizing approaches have assumed activity durations to be independent of each other. Effects of resource sharing and common environmental risk factors on some activities make the assumption of independence to be unrealistic. As a result, we intend to integrate this effect into buffer sizing approaches. We assume that the activities duration of the projects are random variables, then the uncertainty of activities duration can be represented by its variance, and the uncertainty of project duration can be also represented by its variance. By analyzing the impact of dependence between activities on project duration mean and standard deviation, we can draw some illuminations that how much buffer sizes should be set.

3.1 Method of modeling the dependence between activities

The issue of dependence between random variables has been discussed in the areas of project risk analysis (Duffey and Van Dorp, 1998; Van Dorp and Duffey, 1999; Van Dorp, 2005). Duffey and Van Dorp (1998) have developed a model of the statistical dependence for risk-related activities, and consider the dependence between activities as positive dependence. In general, two random variables are positive dependent when large values of one variable tend to be associated with large values of the other random variable. Van Dorp and Duffey (1999) proposed a method to model and quantify positive dependent between uncertainty distributions of activities, with a limitation of considering only one factor in the computational example. Van Dorp (2005) extend the single factor to the multi-factors, and use Monte Carlo simulation to
analyze the effect of dependence between the activity durations on the distribution of the project duration.

In detail, in the method of Van Dorp and Duffey (1999), a directed graph is constructed to represent the dependence diagram between the factors and multivariate (see Figure 1), then the joint distribution between multi-activities affected by the risk factor are transformed into some bivariate joint distributions between the risk factor and the single activity. Every bivariate distribution is modeled by using the method of Copula (Genest and Mackay, 1986), and the Copula is the Diagonal Band distribution first introduced by Cooke and Waij (1986). The degree of dependence from every bivariate distribution is elicited by asking project engineers the percentage of uncertainty in activity duration explained by the risk factor. Finally, a computational example is carried out to draw a conclusion that project duration standard deviation with dependence assumption is more than that using independence assumption. However, in above computational example, the dependence parameter of every bivariate distribution is given, and all the activities in the example project are affected by the same risk factor. Sensitivities of level of dependence parameter and number of risk-related activities on project duration performance aren’t measured, and resource constraints aren’t considered. So the motivation of my paper is to add the two parameters to the formulation of determining buffer sizes that is an important part of critical chain scheduling method.
3.2 Computational example

We intend to analyze this effect by considering a project network with 18 activities (see Figure 2), which is the ninth project in the well-known Patterson’s data set. Both the first activity and the last activity are the dummy activity.

![Figure 2 Activity network of the sample project](image)

3.3 Generation of random durations

In this paper, we assume that the activities duration comes from a right-skewed lognormal distribution used by some other works (Herroelen and Leus, 2001; Tukel et al., 2006). A duration $Y$ is lognormal if $Y = e^X$, with $X$ a normal random variable. If $X$ is normal with mean $\mu$ and standard deviation $\sigma$, then for $Y$ the mean is $\mu_y = e^{\mu+\sigma^2/2}$, and the variance is $\sigma_y^2 = e^{2\mu+\sigma^2} (e^{\sigma^2} - 1) = \mu_y^2 (e^{\sigma^2} - 1)$. We let the mean of $Y$ be equal to the task duration in the Patterson data set, and the standard deviation of every activity duration distribution are given, then the mean and standard deviation for the normal distribution can be calculated. The data of the project are shown in
Table 1. There is one kind of resource, with the available quantities 8.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Duration/days</th>
<th>Standard deviations</th>
<th>Resource</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0.5</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>2.5</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>1.5</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>0.5</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>8</td>
<td>4.5</td>
<td>3</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>13</td>
<td>5</td>
<td>3.5</td>
<td>4</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>0.5</td>
<td>2</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>5</td>
<td>2.5</td>
<td>3</td>
</tr>
<tr>
<td>17</td>
<td>8</td>
<td>3.5</td>
<td>2</td>
</tr>
<tr>
<td>18</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

We intend to use the statistical dependence model developed by Van Dorp and Duffey (1999) to generating random durations with dependence assumption, and the procedures are summarized as follows:

Step 1. Sample the risk factor value $F$ from a Uniform Random Variate on $[0, 1]$.

Step 2. Determine the degree of dependence $\theta$ between the risk factor and every activity.

Ask the project engineers what percentage of every activity duration uncertainty is explained by the risk factor. If the percentage is $K$, then the degree of dependence is $\theta = \sqrt{K}$. $\theta = 0$ implies that the risk factor and the activity are independent, $\theta = 1$ implies that both the variates are identical, $0 < \theta < 1$ denotes
that both the variates are positive dependent.

Step 3. Determine the activity random variate value \( p \) in the Diagonal Band distribution associated with the risk factor and the activity.

1) Sample \( c \) from a Uniform Random Variate on \([0, 1]\);

2) \( a = F - 1 + \theta \), \( b = F + 1 - \theta \);

3) \( P = (b - a) \ast c + a \);

4) if \( P < 0 \), then \( P = -P \), if \( P > 1 \), then \( P = P - 1 \).

Step 4. Generate random durations of the activity \( D \).

Use the inverse of the lognormal cumulative distribution function \( X = \text{logninv}(P, \mu, \sigma) \) in Matlab to generate random durations of the activity. This function returns values at \( P \) of the inverse lognormal cdf with distribution parameters \( \mu \) and \( \sigma \). \( \mu \) and \( \sigma \) are the mean and standard deviation, respectively, of the associated normal distribution. \( P \) is calculated by Sep 3.

Random activities durations with independence assumption are generated by lognormal random numbers function \( R = \text{lognrnd}(\mu, \sigma) \) in Matlab. \( \mu \) and \( \sigma \) are the mean and standard deviation, respectively, of the associated normal distribution.

3.4 Experimental layout

To show the effect of level of dependence parameter and number of risk-related activities on project duration performance, the number of risk-related activities was set for 6, 12, and 18. The level of dependence parameter between the risk factor and every activity are identical, and varied from 0.32 to 0.95 corresponding to the
percentage of every activity duration uncertainty explained by the risk factor varying from 0.1 to 0.9. A thousand replications were generated, and the means and standard deviation of the project were computed for a variety of level of dependence parameter and number of risk-related activities. Thus, the following tables summarize results for a total of $1000 \times 3 \times 9$ runs. For every replication, the risk-related activities were generated randomly. On the other hand, the mean and standard deviation of the project with independence assumption between the activities are also computed as a comparison benchmark.

3.5 Computational results

<table>
<thead>
<tr>
<th>$K$</th>
<th>$\theta$</th>
<th>Number of risk-related activities</th>
<th>Independence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>0.1</td>
<td>0.32</td>
<td>6</td>
<td>30.27</td>
</tr>
<tr>
<td>0.2</td>
<td>0.45</td>
<td>12</td>
<td>30.61</td>
</tr>
<tr>
<td>0.3</td>
<td>0.55</td>
<td>18</td>
<td>30.71</td>
</tr>
<tr>
<td>0.4</td>
<td>0.63</td>
<td></td>
<td>31.10</td>
</tr>
<tr>
<td>0.5</td>
<td>0.71</td>
<td></td>
<td>31.42</td>
</tr>
<tr>
<td>0.6</td>
<td>0.77</td>
<td></td>
<td>31.95</td>
</tr>
<tr>
<td>0.7</td>
<td>0.84</td>
<td></td>
<td>32.66</td>
</tr>
<tr>
<td>0.8</td>
<td>0.89</td>
<td></td>
<td>33.62</td>
</tr>
<tr>
<td>0.9</td>
<td>0.95</td>
<td></td>
<td>35.03</td>
</tr>
</tbody>
</table>

The computational results are summarized in Table 2. The first column $K$ means what percentage of every activity duration uncertainty is explained by the risk factor. The second column $\theta$ means the level of dependence parameter between the risk factor and every activity corresponding to $K$. Mean presents the mean of the project duration. SD presents the standard deviation of the project duration. In the posterior columns, the computational results are given for the four activity categories: the
random 6 activities affected by the risk factor, the random 12 activities affected by the risk factor, all the activities affected by the risk factor and the independence between all the activities.

From Table 2, the results indicate that as the number of risk-related activities increases when $K$ is equal to 0.1 and 0.2, the mean and standard deviation of the project duration don’t increase. Instead, for $K$ equal to 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 and 0.9, as the number of risk-related activities increases, the mean and standard deviation of the project duration increase. The increase in the mean and standard deviation of the project duration is more obvious when $K$ increases. On the other hand, for each given number of risk-related activity, as $K$ and $\theta$ increase both the mean and standard deviation of the project duration with the dependence assumption increase. Especially for all the activities affected by the risk factor, the standard deviation of the project duration for $K$ equal to 0.9 is more than twice as much as that for $K$ equal to 0.1. Compared with the category that all the activities in the project are independent, the mean and standard deviation of the project duration with the dependence assumption are smaller for $K$ equal to 0.1 and 0.2, however, for $K$ equal to the other numerical value the mean and standard deviation of the project duration with the dependence assumption are greater. When all the activities are affected by the risk factor, the standard deviation of the project duration for $K$ equal to 0.9 is near twice as much as that for the independence assumption.

In general, when $K$ is more than or equal to 0.3, the effect of dependence parameter and number of risk-related activities on project duration performance is
valid and positive dependent. This may be relative to the degree of dependence $\theta$ in the statistical dependence model developed by Van Dorp and Duffey (1999). Maybe the model is suitable for $\theta$ that is more than 0.5.

4. Buffer sizing approach with dependence assumption (BSADA)

Based on the computational results in Table 2, under the assumption of the dependence between activities affected by the external risk factor, when the number of risk-related activities is close to the total activity number, it is more likely that delays will occur. For a given number of risk-related activities, as the level of dependence parameter increases, it is again more likely that delays will occur.

So we consider the effect of level of dependence parameter and number of risk-related activities on project duration performance, and a buffer sizing approach with dependence assumption (BSADA) is developed. We assume that there is only one risk factor in the project.

Let $M$ be the number of the activities affected by the risk factor in a feeding chain, $N$ be the total number of activities in this feeding chain, $K_i$ be the percentage of duration uncertainty of activity $i$ explained by the risk factor.

Then for each feeding chain

$$t = \frac{M}{N}$$

$$k = \left(\sum_{i} K_i / M\right)^2$$

$$r = 1 + t \cdot k = 1 + \left(\sum_{i} K_i \right)^2 / (M \cdot N)$$

If project manager want to use the RESM to determining the buffer sizes, and let
\( U_i \) be the duration uncertainty of activity \( i \), the buffer size with the dependence assumption is calculated as:

\[
BufferSize = r \times (\sum_i U_i^2)^{1/2}
\]

When \( M \) is equal to 0, \( r \) is equal to 1, both BSMDA and RSEM are identical. When \( M \) is more than 0, the buffer size generated by the BSADA is more than that generated by the RSEM. When \( M \) and \( K \) are close to the maximum number, the mean and standard deviation of project duration is greater, it is more likely that delays of project will occur. So in order to get the better performance that is the probability of meeting planned completion times, the more buffer size is set under the assumption of the dependence between activities affected by the external risk factor.

5. Conclusions

The most present buffer sizing approaches make the assumption that project activity durations are independent, however, this assumption is unrealistic because of the impact resulted from some external risk factors. In this paper, we analyze the effect of level of dependence parameter and number of risk-related activities on project duration performance, and introduce a method for determining buffer sizes with dependence assumption between activities by means of integrating the two variates to the formulation. Compared with the RSEM, the suggested method can provide better protection when both of the above factor value is at a high level.

Reference


