CHAPTER 4

MODEL FORMULATION USING GRAPHICAL REPRESENTATIONS

4.1 INTRODUCTION

After several years of advancement of commercial computing the system stakeholders still face the problem of accurate effort estimation of software. Software effort estimation is one of the most critical and complex aspect of the software development process because of lack of clear understanding of the calibrating the software data which plays a major role in the effort estimation. Apart from the cost drives and scale factor forecasting the LOC (lines of code), FP (functional points), feature points, object points, model Blitz, wideband Delphi are challenging criteria. Over the last two decades, we have seen a growing trend to use software cost models to estimate the cost of developing software. These models allow us to estimate costs and schedules more quickly and easily than using traditional methods.

A common modification among most models has been to increase the number of input parameters. Some models have been inundated with inputs and output features, yet the accuracy of these models has shown little improvement. Although a great amount of research, time, and money has been devoted to improving our situation, other factors in software development, software complexity, standardization, and lack of data greatly inhibits the ability of software cost model designers to develop credible models. This
leads to a Graphical model which identifies the impact of effort multipliers and includes it accordingly in the process of estimation.

4.2 PROBLEM REPRESENTATION USING GRAPHICAL MODELS

It is observed that the accuracy of the COCOMO II can be viewed as three attributes estimation scheme; the value of estimation variable, the overall value of the effort multipliers and their impact in the estimation scheme. Here the estimation variable is the primary attribute and standard methods available for estimating the estimation variables (e.g. LOC, FP count). It is simple to estimate the overall value of the effort multipliers after assigning the proper values as per the requirements. But the complicated issue is to estimate impact of the effort multipliers, which plays a major role in the estimation scheme and causes for overestimation or underestimation of the software development effort. In this view, this work is aimed at refining the cost drivers and scale factors handling mechanisms in COCOMO II estimation scheme.

There are seventeen effort multipliers which are termed as cost drivers, whose values are qualitatively defined as very low, low, nominal, high, very high, extra high. Based on their impact over the overall effort, these effort multipliers can be devised into two groups; OG and PG.

4.2.1 Optimistic Group (OG)

This group can be defined as a set of effort multipliers of whose range values are inversely proportional to the overall effort to be predicted and it can be described as follows:

OG = \{Effort Multiplier _1, \ Effort Multiplier _2, \ldots, \ Effort Multiplier _9\}
4.2.2 **Pessimistic Group (PG)**

This group can be defined as a set of effort multipliers of whose range values are directly proportional to the overall effort to be predicted and it can be described as follows:

\[ PG = \{ \text{Effort Multiplier}_{10}, \text{Effort Multiplier}_{11}, \ldots, \text{Effort Multiplier}_{17} \} \]

This classification makes a sense in the estimation scheme and plays a vital role in improving the accuracy of the COCOMO II estimation model. The three attributes model can be visualized as a three edged object in a graphical form. In this scheme the overall effort can be estimated in terms of the area(s) of three edged object(s). Let us consider the figure 4.1(a).

In the triangle A, the slope ‘ca1’ represents the overall value of the effort multipliers of the Pessimistic Group (PG). The length of the line ‘ce1’ will be determined accordingly by the value of the estimation variable, here it is LOC measure. By having the lengths of ‘ca1’ and ‘ce1’, the length of ‘œl’ can be determined automatically, which represents the impact of the effort.
multipliers over the overall effort to be estimated. The angle between the lines ‘ca1’ and ‘ce1’ will be determined by the slope of ‘ca1’, which will be determined by overall value of the effort multipliers of the Pessimistic Group (PG). For a constant LOC value, depending on the values of the effort multipliers of the Pessimistic Group (PG), the area under the triangle ‘ca1e1’ will be varied accordingly. Similarly, in the triangle B, the slope ‘cb1’ represents the overall value of the effort multiplier of the Optimistic group (OG). The length of the line ‘cd1’ will be determined accordingly by the value of the estimation variable, here it is LOC measure. By having the length of ‘cb1’ and ‘cd1’, the length of ‘δ1’ can be determined automatically, which represents the impact of the effort multipliers over the overall effort to be estimated. The angle between the lines ‘cb1’ and ‘cd1’ will be determined by the slope of ‘cb1’, which will be determined by overall value of the effort multipliers of the Optimistic Group (OG). For a constant LOC value, depending on the values of the effort multipliers of the Optimistic Group (OG), the area under the triangle ‘cb1d1’ will be varied accordingly.

Let LOC be the height of the triangle, δ1 be the impact of the effort multipliers in the PG set and δ2 be the impact of the effort multipliers in OG set. The effort can be calculated as the areas of the triangles for the PG and OG sets.

The corresponding areas can be calculates respectively as follows using Equation (4.1, 4.2).

\[ \text{PG } E_{2a} = 0.5 \times \delta_1 \times \text{height} \]  \hspace{1cm} (4.1)

\[ \text{OG } E_{2a} = 0.5 \times \delta_2 \times \text{height} \]  \hspace{1cm} (4.2)

where,

\text{PG } E_{2a} \text{ is the effort due to the Pessimistic Group in the Figure 4.1(a)}

and \text{OG } E_{2a} \text{ is the effort due to the Optimistic Group in the Figure 4.1(a).}
Figure 4.1 Model Formulation using Graphical Representations
Now based on the insights of the deductive and additive theories of impact of effort multipliers over the overall effort, the values of $\delta_1$ and $\delta_2$ can be assumed accordingly.

By following the deductive principle, the value of $\delta_1$ can be taken as $\text{LOC}/\text{EM}$ and by following the additive principle, the value of $\delta_2$ can be taken as $\text{LOC} \times \text{EM}$. This cognitive assumption improves the accuracy of the COCOMO estimation scheme to a larger extent using graphical model.

By substituting the values of $\delta_1$ and $\delta_2$, the Equations (4.1) and (4.2) can be rewritten as follows:

$$
\begin{align*}
^P \text{G}E_{2a} &= 0.5 \times \text{LOC} \times \text{EM} \times \text{LOC} \\
^O \text{G}E_{2a} &= 0.5 \times \text{LOC} / \text{EM} \times \text{LOC}
\end{align*}
$$

(4.3)

Then the total effort is defined as the product of efforts due to OG and PG and it can be calculated as follows:

$$
E_{2a} = ^P \text{G}E_{2a} \times ^O \text{G}E_{2a}
$$

(4.4)

Hence the overall effort includes the corresponding impacts of both the types of effort multipliers.

### 4.3 VARIOUS RANGES OF LOC AND EM

An important and interesting point to be noted in this scheme is, though by keeping the values of effort multipliers as constants, the corresponding impact may be varied in accordance to the variations in the value of LOC measure. But in the earlier estimation scheme, in such
situations, the variations in the overall effort is the function of corresponding variations in the LOC measure alone and the impact of the effort multipliers will not be taken into the account. This interesting property of this scheme can be explained in detail under various conditions as follows:

4.3.1 Case 1- Increase in LOC

It is already discussed in the chapter 2.4 that the value of LOC has the direct impact on the overall effort. i.e. overall effort increases along with the increments in the LOC value and the overall effort decreases along with the decrements in the LOC value by keeping the other parameters as constant. In the figure 4.1(b), the LOC value is increased from LOC to LOC1 and these increments are clearly represented in both the triangles. In the figure 4.1(b), the variations are done only based on the LOC values, but these variations are automatically reflected in the impacts of the effort multipliers. i.e. the value of $\delta 1$ is increased to $\delta 4$ and the value of $\delta 2$ is increased to $\delta 3$. Since the overall effort is the function of areas of the triangles, the increments in the values of $\delta 1$ and $\delta 2$ (as $\delta 4$ and $\delta 3$) will be correspondingly reflected in the overall effort. From the Figure 4.1(b), the corresponding overall effort can be estimated as follows:

$$E_{2b} = ^{PG}E_{2b} \times ^{OG}E_{2b}$$  \hspace{1cm} (4.5)

where,

$E_{2b}$ is the overall effort from the figure 4.1(b), $^{PG}E_{2b}$ is the effort due to the Pessimistic Group in the figure 4.1(b) and $^{OG}E_{2b}$ is the effort due to the Optimistic Group in the Figure 4.1(b).

The values of $^{PG}E_{2b}$ and $^{OG}E_{2b}$ can be calculated as follows:
The Equation (4.6) can also be written as follows

\[
^{PG}E_{2b} = 0.5 \delta_4 \cdot \text{height}
\]  

(4.6)

\[
^{OG}E_{2b} = 0.5 \delta_5 \cdot \text{height}
\]  

(4.7)

\[
^{OG}E_{2b} = 0.5 \frac{LOC_1}{EM} \cdot \text{LOC}_1
\]

The Equation (4.6) can also be written as follows

\[
^{PG}E_{2b} = 0.5 \left( \delta_1 + \Delta \delta_{PG} \right) \cdot (LOC + \Delta LOC)
\]  

(4.8)

where,

\[
\Delta \delta_{PG} \text{ is the change in the impact of the effort multiplier of PG and it can be calculated as } \delta_4 - \delta_1
\]

\[
\Delta \text{LOC is the change in the LOC measure and it can be calculated as LOC}_1 - \text{LOC}.
\]

Similarly, the Equation (4.7) can also be written as

\[
^{OG}E_{2b} = 0.5 \left( \delta_2 + \Delta \delta_{OG} \right) \cdot (LOC - \Delta LOC)
\]  

(4.9)

where,

\[
\Delta \delta_{OG} \text{ is the change in the impact of the effort multiplier of OG and it can be calculated as } \delta_3 - \delta_2
\]

\[
\Delta \text{LOC is the change in the LOC measure and it can be calculated as LOC}_1 - \text{LOC}.
\]
Hence, in this case, the increment is done in the LOC measure only. But, though there is no change in the values of effort multipliers, the corresponding impact has been taken into account in accordance to the variation in the value of LOC and accordingly the overall has been estimated. This significant property can be realized through the equations (4.8) and (4.9) as in the above case.

4.3.2 Case 2 – Decrease in LOC

As it is already pointed in chapter 2.4 that the value of LOC has the direct impact on the overall effort. i.e. overall effort increases along with the increments in the LOC value and the overall effort decreases along with the decrements in the LOC value by keeping the other parameters as constant. In the figure 4.1(c), the LOC value is decreased from LOC to LOC1 and these decrements are clearly represented in both the triangles. In the figure 4.1(c), the variations are done only based on the LOC values, but these variations are automatically reflected in the impacts of the effort multipliers. i.e. the value of \( \delta 1 \) is decreased to \( \delta 4 \) and the value of \( \delta 2 \) is decreased to \( \delta 3 \). Since the overall effort is the function of area of the triangles, the decrements in the values of \( \delta 1 \) and \( \delta 2 \) (as \( \delta 4 \) and \( \delta 3 \)) will be correspondingly reflected in the overall effort. From the Figure 4.1(c), the corresponding overall effort can be estimated as follows:

\[
E_{2c} = ^{PG}E_{2c} \ast ^{OG}E_{2c}
\]

(4.10)

where,

\( E_{2c} \) is the overall effort from the Figure 4.1(c), \(^{PG}E_{2c}\) is the effort due to the Pessimistic Group in the Figure 4.1(c) and \(^{OG}E_{2c}\) is the effort due to the Optimistic Group in the Figure 4.1(c).
The values of \(^{PG}E_{2c}\) and \(^{OG}E_{2c}\) can be calculated as follows:

\[ {^{PG}E}_{2c} = 0.5 \cdot \delta_4 \cdot \text{height} \]

\[ {^{PG}E}_{2c} = 0.5 \cdot LOC_1 \cdot EM \cdot LOC_1 \quad (4.11) \]

\[ {^{OG}E}_{2c} = 0.5 \cdot \delta_3 \cdot \text{height} \]

\[ {^{OG}E}_{2c} = 0.5 \cdot LOC_1 \cdot EM \cdot LOC_1 \quad (4.12) \]

The Equation (4.11) can also be written as follows:

\[ {^{PG}E}_{2c} = 0.5 \cdot (\delta_1 - \Delta \delta_{PG}) \cdot (LOC - \Delta LOC) \quad (4.13) \]

where,

\( \Delta \delta_{PG} \) is the change in the impact of the effort multiplier of PG and it can be calculated as \( \delta_1 - \delta_4 \).

\( \Delta LOC \) is the change in the LOC measure and it can be calculated as \( LOC - LOC_1 \).

Similarly, the Equation (4.12) can also be written as

\[ {^{OG}E}_{2c} = 0.5 \cdot (\delta_2 - \Delta \delta_{OG}) \cdot (LOC - \Delta LOC) \quad (4.14) \]

where,

\( \Delta \delta_{OG} \) is the change in the impact of the effort multiplier of OG and it can be calculated as \( \delta_2 - \delta_3 \).
\( \Delta \text{LOC} \) is the change in the LOC measure and it can be calculated as \( \text{LOC} - \text{LOC1} \).

Hence, in this case, the decrements are done in the LOC measure only. But, though there is no change in the values of effort multipliers, the corresponding impact has been taken into account in accordance to the variation in the value of LOC and accordingly the overall has been estimated. This significant property can be realized through the equations (4.13) and (4.14) as in the above case.

### 4.3.3 Case 3 – Increase in EM

As mentioned earlier in chapter 2.4 that the range of EM also has the impact on the overall effort. i.e. overall effort increases along with the increments in the range of PG set and the overall effort decreases along with the increments in the range of OG set by keeping the other parameters as constant. In the figure 4.1(d), the EM range of both the groups are increased. When the range of OG group increased from EM to EM1 then the slope of the line \( cb1 \) increases and falls to the point \( b2 \), i.e \( \delta 2 \) to \( \delta 3 \), thus the effort decreases in OG group, similarly for PG group if the range increases from EM to EM1, then the slope of the line \( ca1 \) increases to point \( a2 \) as shown in the figure 4.1(d). In the figure 4.1(d), the variations are done only based on the EM values, but these variations are automatically reflected in the impacts of the effort multipliers. i.e. the value of \( \delta 1 \) is increased to \( \delta 4 \) and the value of \( \delta 2 \) is decreased to \( \delta 3 \). Since the overall effort is the function of area of the triangles, the change in the values of \( \delta 1 \) and \( \delta 2 \) (as \( \delta 4 \) and \( \delta 3 \)) will be correspondingly reflected in the overall effort. From the figure 4.1(d), the corresponding overall effort can be estimated as follows:

\[
E_{2d} = \frac{P_G}{E_{2d}^{OG}} E_{2d}^{OG} \]

\[(4.15)\]
where,

\[ E_{2d} \] is the overall effort from the figure 4.1(d), \( {^PG}E_{2d} \) is the effort due to the Pessimistic Group in the figure 4.1(d) and \( {^OG}E_{2d} \) is the effort due to the Optimistic Group in the Figure 4.1(d).

The values of \( {^PG}E_{2d} \) and \( {^OG}E_{2d} \) can be calculated as follows:

\[
{^PG}E_{2d} = 0.5 \cdot \delta_4 \cdot \text{height} \\
{^PG}E_{2d} = 0.5 \cdot \text{LOC} \cdot \text{EM}_1 \cdot \text{LOC} \\
{^OG}E_{2d} = 0.5 \cdot \delta_3 \cdot \text{height} \\
{^OG}E_{2d} = 0.5 \cdot \text{LOC} / \text{EM}_1 \cdot \text{LOC}
\] (4.16) (4.17)

The Equation (4.16) can also be written as follows

\[
{^PG}E_{2d} = 0.5 \cdot (\delta_1 + \Delta \delta_{PG}) \cdot \text{LOC} \\
{^PG}E_{2d} = 0.5 \cdot (\delta_1 + \Delta \delta_{PG}) \cdot \text{LOC}
\] (4.18)

where,

\[ \Delta \delta_{PG} \] is the change in the impact of the effort multiplier of PG and it can be calculated as \( \delta_4 - \delta_1 \).

Similarly, the Equation (4.17) can also be written as

\[
{^OG}E_{2d} = 0.5 \cdot (\delta_2 - \Delta \delta_{OG}) \cdot \text{LOC} \\
{^OG}E_{2d} = 0.5 \cdot (\delta_2 - \Delta \delta_{OG}) \cdot \text{LOC}
\] (4.19)
where,

\[ \Delta \hat{\delta}_{OG} \] is the change in the impact of the effort multiplier of OG and it can be calculated as \( \delta_2 - \delta_3 \).

Hence, in this case, the increments are done in the EM measure only. But, though there is no change in the values of LOC, the corresponding impact has been taken into account in accordance to the variation in the value of EM and accordingly the overall has been estimated. This significant property can be realized through the equations (4.18) and (4.19) as in the above case.

4.3.4 Case 4 – Decrease in EM

It is already discussed in chapter 2.4 that the range of EM also has the impact on the overall effort. i.e. overall effort increases along with the increments in the range of PG set and the overall effort decreases along with the increments in the range of OG set by keeping the other parameters as constant. In the figure 4.1(e), the EM range of both the groups is decreased. When the range of OG group decreased from EM to EM1 then the slope of the line cb1 decreases and falls to the point b2, i.e \( \delta_2 \) to \( \delta_3 \), thus the effort increases in OG group, similarly for PG group if the range decreases from EM to EM1, then the slope of the line ca1 decreases to point a2 as shown in the figure 4.1(e). In the figure 4.1(e), the variations are done only based on the EM values, but these variations are automatically reflected in the impacts of the effort multipliers. i.e. the value of \( \delta_1 \) is decreases to \( \delta_4 \) and the value of \( \delta_2 \) is increased to \( \delta_3 \). Since the overall effort is the function of area of the triangles, the change in the values of \( \delta_1 \) and \( \delta_2 \) (as \( \delta_4 \) and \( \delta_3 \)) will be correspondingly reflected in the overall effort. From the Figure 4.1(e), the corresponding overall effort can be estimated as follows:
\[ E_{2e} = PGE_{2e} \times OG E_{2e} \]  \hspace{1cm} (4.20)

where,

\( E_{2e} \) is the overall effort from the Figure 4.1(e), \( PGE_{2e} \) is the effort due to the Pessimistic Group in the Figure 4.1(e) and \( OG E_{2e} \) is the effort due to the Optimistic Group in the Figure 4.1(e).

The values of \( PGE_{2e} \) and \( OG E_{2e} \) can be calculated as follows:

\[ PGE_{2e} = 0.5 \times \delta_4 \times \text{height} \]

\[ PGE_{2e} = 0.5 \times LOC \times EM_1 \times LOC \] \hspace{1cm} (4.21)

\[ OG E_{2e} = 0.5 \times \delta_3 \times \text{height} \]

\[ OG E_{2e} = 0.5 \times LOC / EM_1 \times LOC \] \hspace{1cm} (4.22)

The Equation (4.21) can also be written as follows:

\[ PGE_{2e} = 0.5 \times (\delta_4 - \Delta \delta_{PG}) \times LOC \] \hspace{1cm} (4.23)

where,

\( \Delta \delta_{PG} \) is the change in the impact of the effort multiplier of PG and it can be calculated as \( \delta_1 - \delta_4 \).

Similarly, the Equation (4.22) can also be written as

\[ OG E_{2e} = 0.5 \times (\delta_3 + \Delta \delta_{OG}) \times LOC \] \hspace{1cm} (4.24)
where,

\[ \Delta \delta_{OG} \] is the change in the impact of the effort multiplier of OG and it can be calculated as \( \delta 3 - \delta 2 \).

Hence, in this case, the decrements are done in the EM measure only. But, though there is no change in the values of LOC, the corresponding impact has been taken into account in accordance to the variation in the value of EM and accordingly the overall has been estimated. This significant property can be realized through the equations (4.23) and (4.24) as in the above case.

### 4.3.5 Variations of EM in OG and PG

It is already discussed in chapter 2.4 that the range of EM also has the impact on the overall effort. i.e. overall effort increases along with the increments in the range of PG set and the overall effort decreases along with the increments in the range of OG set by keeping the other parameters as constant. In the figure 4.1(f), the EM range of OG group is decreased and the EM range of PG group is increases. When the range of OG group decreased from EM to EM1 then the slope of the line cb1 decreases and falls to the point b2, i.e \( \delta 2 \) to \( \delta 3 \), thus the effort increases in OG group, similarly for PG group if the range increases from EM to EM1, then the slope of the line ca1 increases to point a2 as shown in the figure 4.1(f). In the figure 4.1(f), the variations are done only based on the EM values, but these variations are automatically reflected in the impacts of the effort multipliers. i.e. the value of \( \delta 1 \) is increases to \( \delta 4 \) and the value of \( \delta 2 \) is increased to \( \delta 3 \).

Since the overall effort is the function of area of the triangles, the change in the values of \( \delta 1 \) and \( \delta 2 \) (as \( \delta 4 \) and \( \delta 3 \)) will be correspondingly
reflected in the overall effort. From the Figure 4.1(f), the corresponding overall effort can be estimated as follows:

\[
E_{2f} = \frac{PG}{E_{2f}} \times \frac{OG}{E_{2f}}
\]

(4.25)

where,

\(E_{2f}\) is the overall effort from the Figure 4.1(f), \(PG\)\(E_{2f}\) is the effort due to the Pessimistic Group in the Figure 4.1(f) and \(OG\)\(E_{2f}\) is the effort due to the Optimistic Group in the Figure 4.1(f).

The values of \(PG\)\(E_{2f}\) and \(OG\)\(E_{2f}\) can be calculated as follows:

\[
PG\ E_{2f} = 0.5 \times \delta_4 \times \text{height}
\]

(4.26)

\[
PG\ E_{2f} = 0.5 \times LOC \times EM_1 \times LOC
\]

\[
OG\ E_{2f} = 0.5 \times \delta_3 \times \text{height}
\]

(4.27)

\[
OG\ E_{2f} = 0.5 \times LOC / EM_1 \times LOC
\]

The Equation (4.26) can also be written as follows

\[
PG\ E_{2f} = 0.5 \times (\delta_l + \Delta \delta_{PG}) \times LOC
\]

(4.28)

where,

\(\Delta \delta_{PG}\) is the change in the impact of the effort multiplier of PG and it can be calculated as \(\delta_4 - \delta_1\).
Similarly, the equation (4.27) can also be written as

\[ ^{OG}E_{2j} = 0.5 \times (\hat{\delta}_2 + \Delta \hat{\delta}_{OG}) \times LOC \]

(4.29)

where,

\[ \Delta \hat{\delta}_{OG} \] is the change in the impact of the effort multiplier of OG and it can be calculated as \( \hat{\delta}_3 - \hat{\delta}_2 \).

Hence, in this case, the changes are done in the EM measure only. But, though there is no change in the values of LOC, the corresponding impact has been taken into account in accordance to the variation in the value of EM and accordingly the overall has been estimated. This significant property can be realized through the equations (4.28) and (4.29) as in the above case.

By using the Equations (4.3) and (4.4) the overall effort can be estimated. Since there is no appropriate method available for estimating the impact of the effort multipliers, this scheme may be the first to estimate the impact of the effort multipliers as a function of LOC. This scheme yields a considerable degree of underestimation in overall effort estimation.

The function \( F(x) \) is given as an input for GA. This does minimization on the \( f(x) \) for the offsprings by finding the value of ‘x’

\[ f(x) = \left( \prod_{i=1}^{\hat{\beta}} 0.5 \times LOC^2 \times ^{PG} EM_i \right) \prod_{j=1}^{\hat{\rho}} 0.5 \times LOC^2 / ^{OG} EM_j / LOC \]

(4.30)

Where \( 0 < x < 1 \) and \( LOC \geq 1 \)
4.4 SUMMARY

It is observed that the accuracy of the COCOMO II can be viewed as three attributes estimation scheme; the value of estimation variable, the overall value of the effort multipliers and their impact in the estimation scheme. Here the estimation variable is the primary attribute and standard methods available for estimating the estimation variables (e.g. LOC, FP count). It is simple to estimate the overall value of the effort multipliers after assigning the proper values as per the requirements. The seventeen effort multipliers of COCOMO II are categorized as Optimistic Group and Pessimistic Group and the values of these range from very low, low, nominal, high, very high, extra high. This research work is aimed at refining the cost drivers and scale factors handling mechanisms in COCOMO II estimation scheme using graphical representation.