A Systematic Mapping Study on Value of Reuse

Mandava Kranthi K          B M Konda          Dr. K. Thammi Reddy          B. Ravi Kiran
Asst. Prof, Dept. of CSE          Student          Professor, Dept of CSE          Sr.Asst.Prof, Dept of CSE
ANITS, Vizag, India              BTH, Sweden                  GITAM, Vizag, India                  ANITS, Vizag, India

A.Vindhya
Student
ANITS, Vizag, India

ABSTRACT
Software reuse is considered as the key to a successful software development because of its potential to reduce the time to market, increase quality and reduce costs. This increase in demand made the software organizations to envision the use of software reusable assets which can also help in solving reoccurring problems for successful software. Now a day, organizations are interested in implementing reuse program. As the “reuse” is growing in software industry, there is a growing need to assess the value of reuse by measuring it, which helps to know their success. As the concepts like reuse and reusability emerged, a question arose on how to measure them. So, in our paper, we investigate on what techniques, methods, models and metrics for assessing the value of reuse and proposed some new subcategories for the reuse metrics and models categories.

Note: This work is part of our thesis which we have submitted during our masters in BTH, Sweden. One can find the whole thesis at [55].

General terms
Software Engineering, Software Reuse

Keywords
Software, Reuse, Reusable Assets, Value, Metrics, Models.

1. INTRODUCTION
Software reuse is the process of creating a new system from that of existing system rather than creating the new one from the scratch. In other words, it is the reusing of existing software artifacts or software assets to build a new system. The concept of software reuse has been introduced to overcome the software crisis i.e., the problem of building large and reliable software system in a cost effective and controlled way by McIlroy in 1968 [48]. Initially, software reuse was limited to source code. Due to the increase in customer needs and market demand for sophisticated software, the software companies started thinking beyond source code. This leads to the reuse of other life cycle assets. By reusing the other life cycle assets like design, algorithms, knowledge etc, and new software can be brought into the market faster. Software reuse helps in not only reducing the time but also reduces the cost [49] [48].

As the “reuse” is growing in software industry, there is a growing need to assess the value of reuse by measuring it, which helps to know their success. As the concepts like reuse and reusability emerged, a question arose on how to measure them in order to get success through reuse. For measuring it, reuse metrics and models have been defined for many areas of software reuse and
categorized into 6 categories [1]. Assessing the value of reuse is a major concern in the software industry. For assessing its value, reuse should be measured by using the metrics and models. Measuring reuse will help the organization to know their progress in software reuse, to know how much amount of reuse is done or to assess the cost benefits of software reuse etc. For this, W. B. Frakes in 1996 has done a review on some of the existing important models or metrics or methods. However, there are no widely accepted models and the organizations are still unsure of getting success by using those models which are predicted [50]. For measuring reuse, reuse metrics and models have been defined for many areas of software reuse and categorized into 6 categories in [1] as Cost Benefit Analysis Models Maturity Assessment Models, Amount of Reuse Metrics, Failure Modes Model, Reusability Assessment, Reuse Library Metrics.

2. TAXONOMY OF REUSE METRICS AND MODELS
We present taxonomy of reuse metrics and models in which different categories and sub-categories of reuse metrics/models/methods are presented. This taxonomy is based on the taxonomy defined by Frakes in [1]. Frakes [1] in his taxonomy does not show subcategories. But going deep into the report, we could find that some categories do have the subcategories. And to his taxonomy, we have added some other subcategories in cost benefit analysis models and amount of reuse metrics. These are not mentioned by Frakes [1]. But, we have gone through other studies of Jorge Mascena [5], Frakes [35] and Suri [4] in which they have mentioned the subcategories to amount of reuse metrics category along with those mentioned by Frakes [1].

The reuse metrics and models are divided into 6 categories as shown figure 1 [1, 2, and 55].

1. Cost Benefit Analysis Models
2. Maturity Assessment Models
3. Amount of Reuse Metrics
4. Failure Modes Model
5. Reusability Assessment
6. Reuse Library Metrics
2.1 Cost Benefits Analysis Models
Cost benefit analysis helps to know the cost benefits of implementing reuse. These models include economic cost benefit analysis, return on investment, quality of investment and productivity pay-offs. These models are for assisting the organization in estimating their cost, effort, and time which is involved in systematic reuse.

“The value of software reuse refers to whether it is more cost effective, in terms of time, money, or personnel, to reuse software as opposed to developing it from scratch each time it is needed” Frakes [6].

2.1.1 Economic Cost Benefit Analysis:
Economic Cost Benefit Analysis helps in assessing the costs of reusing a reusable component.

2.1.2 Return on Investment Analysis:
It is one of several approaches to evaluating and comparing investments. It helps us to know the benefits. A good Return on Investment means that the investment returns compare favorably to investment costs. This analysis is crucial for reuse investments.

2.1.3 Quality of Investment:
Quality of investment helps in making a good reuse investment.

2.1.4 Business Reuse Metrics:
These metrics help in assessing or estimating the effort saved by reuse.

2.2 Maturity Assessment Models
Maturity assessment is needed by an organization in assessing the degree of maturity of its reuse implementation process. Reuse maturity assessment models will help the organizations in estimating how advanced the reuse programs are in implementing systematic reuse [1]. This helps the organizations to know their progress in implementing reuse programs.

2.3 Amount of Reuse Metrics
Amount of reuse metrics is used to estimate how much of reuse is done in a give life cycle object. According to [1], amount of reuse metrics are used to assessing and also monitoring the reuse improvement effort by tracking of the percentages of reuse for life cycle objects. The amount of reuse metrics is subdivided into six types:

2.3.1 Reuse level
Reuse level is the ratio of number of reused items to the total number of items [7, 9].

2.3.2 Reuse percent
Reuse percent is the ratio of number of reused lines of code to the total number of lines of code [8, 9].

2.3.3 Reuse frequency
Reuse frequency is the ratio of number of references to the reused items to the total number of references [7, 9].

2.3.4 Reuse ratio:
Reuse ratio considers partially changed items as reused and is same as reuse percent [9, 10].

2.3.5 Reuse density:
Reuse density is the ratio of number of reused parts to the total number of lines of code [9].

2.3.6 Reuse size and frequency:
Reuse size and frequency is similar to reuse frequency and considers the size of items in number of lines of code (LOC) [9, 10].

2.4. Failure Modes Models
Failure modes analysis is used to identify and order the obstacles to reuse in an organization. Failure modes analysis gives us an approach for measuring the reuse process and improving it which is based on a model of the ways a reuse process fails [1, 11].

2.5. Reusability Assessment
Reusability metrics indicate the possibility that an artifact is reusable or the readiness of an artifact or asset to be reusable. In this the attributes of a component which indicate its reusability are measured [1].

2.6. Reuse Library Metrics
Reuse library metrics are used for managing and tracking the reuse repository usage. The Indexing schemes in the reuse library are evaluated by using these metrics. For evaluating the indexing schemes the reuse library metrics are [1]:

2.6.1 Indexing costs:
Measuring the cost of creating, maintaining, updating a classification scheme.

2.6.2 Searching effectiveness:
Assess how well the classification schemes help users to search effectively for reusable components.

2.6.3 Support for understanding:
Measures how well a classification scheme helps the users to understand the components.

2.6.4 Efficiency:
Measure the efficiency of reusable library in terms of memory, fastness etc.

In addition to this, Quality of the assets is also a measure for reuse library metrics which was derived by Frakes in 1987.

3. ANALYSIS
3.1 Bubble graph
The size of the bubble depends upon the number of studies in that bubble. The bubbles at the intersection of the axes contain reference numbers of the studies. The X-Axis is divided in to two halves i.e., the left and right halves. On the right half of the X-Axis in figure 2, we show the validation status of the studies and also indicate which type of validation; the study falls in to (like industrial case study, academic case study, academic experiment, industrial experiment, survey). On the left half of the X-Axis we present the studies which proposed a method, model, metrics or an approach for measuring reuse to assess its value. The Y-axis has six reuse metrics and models categories (Cost benefit analysis models, Maturity assessment models, Amount of reuse metrics, Failure modes models, Reusability assessment, and Reuse library metrics).

We have performed our investigation on the research articles between the years 1968 to 2010. This research question is answered using the populated taxonomy of W B Frakes [1]. The remaining research regarding this research question is based on this taxonomy. Most of the research regarding measuring the reuse to assess its value is done from 1987. As initially the reuse is considered only for coding the earlier authors discussed on methods/models/metrics for assessing the value of reuse keeping in mind, only the coding part. But later as per the observation of the recent research studies, the researchers seemed to have
understood the importance of introducing the reuse to other life cycle objects (reusable assets like requirements, design etc) and they have extended this concept from coding to other life cycle objects.

From 1987 there are many methods/models/metrics discussed or presented regarding measuring the reuse to assess its value. All these are divided into different categories based on their application to different areas of software reuse and applied to the Frakes [1] taxonomy. There are six categories as presented in section 2.

1. Cost Benefit Models
2. Maturity Assessment Models
3. Amount of Reuse Metrics
4. Failure Modes Models
5. Reusability Assessment Models
6. Reuse Library Metrics

### 3.2. State of Validation

We present a graph (in figure 3) to show the validation status of each reusable asset. In the graph, X-Axis represents the reusable assets and Y-Axis represents the percentage of validation (i.e., number of validated and non-validated studies and reviews as well). As the gathered studies also contain reviews which don’t come under validated or non-validated studies, they are presented in the graph along with the validated and non-validated studies. The validated and non-validated studies along with reviews will sum up to 100 percent.

#### 3.2.1 Overall Validation Status and Analysis

1. Though the research in this area started from 1968, we have mainly focused the research studies in between 1987 to 2009. According to the observation there is less initiation to validate the existing models or maybe they have been validated by the organizations but were not reported.
2. Among the found studies 86% are research studies on metrics or models or methods that we have found during our systematic review, in those only 36% of them are validated and 48% of them are non-validated and remaining 16% are review studies. Based on the observation non-validated models are a bit higher than the validated.
3. Among the validated metrics/models/methods, 38.8% are industrially validated and 55.5% are academically validated and 5.5% are validated through surveys.
4. Observation shows that, not many industries actually put effort to report their experiences in using a particular metric/method/model. If they had reported the experiences, it would be easier for other organizations to decide whether to use a particular metric or not based on the experience that is reported.
5. Some efforts are kept to validate academically, but that too not many.
6. Based on the observations not many studies have actually presented the limitations as many had just proposed a model/method/metric without any actual validation and many authors tried to present the advantages of their model. But some authors tried to find the limitations in the previous models like for example Rine in their study work [106] reuse capability model of study [85] proved to be unstable.
7. 3 studies have tried to validate the metrics/methods/models that are just proposed in the previous studies.
8. Many studies have validated their models without using the industrial data but the validated by setting up random values.
9. Review studies also played a key role in this field of research. Around 17% of the studies that found were reviews. Some tried to review the previous studies and kept efforts to validate the models/metrics/methods and some reviewed the previous studies to find out the trends and recent happenings in this field of research like the study of Parastoo Mohagheghi [17] which reviewed the studies from 1994-2005 to gather the evidences for successful software reuse programs in the industry and explained several drawbacks in the present studies and reported that the most industrial studies are of the observational type. Researchers used data that is collected in the industrial settings but will have little control over the environment or the data collection procedures and [2] by Frakes, which presented the present status and the future of software reuse. In 1996 Frakes [1] reviewed surveyed some important reuse metrics and models and defined a taxonomy of reuse metrics and models and published his observations. These reviews are beneficial to any researcher who wants to know the recent activities in this field of research.

#### 3.2.2 Validation status for each Category

Here in this section, we present validation status of each Reuse metrics and models category in detail, along with the percentage of review studies and the gap we have found and our suggestions to fill the gap.

#### 3.2.2.1 Cost Benefit Models

From the total number of hits for this category, we have selected 15 studies using inclusion and exclusion criteria. Among those, 10 are validated and 2 are non validated model or metric or method, 3 are review studies of the total 15 studies and among

![Figure 3: Validation Status- Value of Reuse](image-url)
3.2.2.2. Maturity Assessment Models
We have selected 9 studies regarding this study using inclusion and exclusion criteria. In those, 1 study is validated and 7 are non-validated models/metrics/methods and 1 is a review paper. Among those 3 studies are extensions to the previous ones, 1 study is related to the validation of one of the 9 studies and Study [23] is the extension to [22], study [24] is extension to study [23] and study [26] is the extension to study [24]. And study [25] is related to the validation of study [24]. 1 study [23] is related to STARS project and 1 study [28] is related to RiSE project. These models help the organization to assess the advancement of the reuse programs in implementing the systematic reuse. Though very interesting models were reported around 77.7% of the models in the studies that we have found are non-validated, only 11.1% of the models are validated and 11.2% are reviews. The only one validated study is only validated through survey and there are no reports with industrial validation. So, much more validation process is needed for the non-validated models and if this validation is done in industrial scenario or industrial environment, then it will be useful for the organizations to choose the model (using the validation results) that best suits for it. All the validated and non-validated models had the common aim of measuring the maturity of the reuse program in implementing the systematic reuse. From 2004 onwards, not much research is done in this field.

3.2.2.3. Amount of Reuse Metrics
We have chosen 15 studies regarding this study. In those 3 studies are validated and 6 are non-validated models/metrics/methods, 6 are review studies and study [31] is an extension to study [30] and there is not much validation is done in the past studies regarding this amount of reuse metrics study.

The amount of reuse metrics is subdivided into six categories as discussed in section 3.2.3. Almost all the studies discussed about the general or basic amount of reuse metrics, reuse level and reuse frequency and a very few articles like Poulin [21], Frakes [7], Basili [36], Devanbu [10], Frakes [35], Suri [4], Mascena [5, 9] were found which discussed on other reuse metrics like reuse percentage, reuse size and frequency, reuse ratio, reuse rate. Among the found studies around 20% are validated, 40% of the studies reported are non-validated models/metrics/methods and another 40% of studies are reviews. From the percentage figures, we can observe that there are more non-validated and review studies regarding amount of reuse metrics and so, much more effort should be kept in validating the models mainly in the industrial scenario.

3.2.2.4. Failure Modes Models
We could find 2 studies that could best suit for our cause and study [11] is a non-validated model and study [1] is a review study.

W Frakes in [1] [11] was the only author who presented and mentioned about the failure modes model in his studies, according to our observation. The concept is very useful for the organization to make them know the obstacles for reuse. So it is good to have much more research in this field is required.

3.2.2.5. Reusability Assessment Models
We could find 9 studies important regarding this study. In those 3 are validated and 5 are non-validated models or metrics or methods or frameworks. 1 study is a review study.

The 8 studies that are found regarding reusability assessment had a common aim of assessing the readiness of an artifact to be reusable or to indicate the possibility that an artifact is reusable. Only 33.3% of the models are validated and 55.5% of models are non-validated and 11.1% is review study. The research in this field from 1997 to 2003 seems to be not much. We strongly recommend much research to be done in this area focusing on the other reusable assets (which are other than coding) by not sticking to code itself.

3.2.2.6. Reuse Library Metrics
For each type of reuse library metrics we have searched and we could find 5 studies regarding this study. Only 1 study is validated and there are 4 non-validated models or metrics. We could find more non-validated studies than that of validated. There is very less research in this category. The studies that were found have the common aim helping in managing and tracking the reuse repository usage. We could find 7 models of
which only 20% are validated and other 80% are non-validated studies.
In figure 4, we have shown the percentage of academic, industrial and survey validations under each category. We can notice that industrial validations are very less with 39% when compared to academic validation with 56%. When considering the cost benefit analysis we could notice that academic and industrial validations are equal in ratio. Only one study used survey for the validation purpose. Since, the survey is conducted in an industrial environment; it can be treated as an industrial validation. The figure 4 shows that the academic validations are more than the industrial validations.

Figure 4: Percentage of Academic, Industrial and Survey Validations

3.3 Areas in focus
In this section, we present a surface graph (in figure 5) which shows the assets in focus. Figure 5 shows the number of studies found per year.
Cost benefit analysis, Maturity assessment model, Amount of reuse metrics were focused more than the Failure modes models, Reuse library metrics and Reusability assessment. As per the observation of studies from 1987 to 2009 Cost benefit analysis, Maturity assessment model, Amount of reuse metrics and Reusability assessment were equally focused by the researchers. There are studies regarding these which were published in close intervals of time. But on observing the Failure modes models, it was mainly mentioned by W.B Frakes [1] [11] both in year 1996 and there is no recent publication regarding this category. Regarding reuse library metrics, some research was done and upon observing the studies from 1987 to 2009, the gap between each published study is much greater for example one study published in 1987 and next important study published in 1994. On coming to reusability assessment there is a large gap in between the year 1997 and 2003.

Figure 5: Areas in Focus
Through this graph, we can identify which areas are in focus for a particular year. For example there was no contribution in the year 2001 for all six categories. We can also notice that, the research contribution is more during the year 2006. There is no contribution on cost benefit models during the period between 1997 and 2001. The contribution on the failure mode models is only present between 1995 and 1997.

3.3.1 Representations methods for methods / models / metrics each category
Here we present the percentage of representation methods used by the authors to represent their methods/models/metrics etc;
3.3.1.1. Cost Benefit Analysis Models
Systematic review results shows that approximately 77.7% of the studies represented their metrics/models/methods through the mathematical means and about 22.3% of the studies represented through diagram or tables or theories. Other studies used the diagrams or table or theory to represent or describe their metrics/models/methods.
3.3.1.2. Maturity Assessment Models
Systematic review results shows that approximately 16.6% of the studies represented their metrics/models/methods through the matheatical means and about 83.4% of the studies represented their metrics/models/methods through diagrams or tables or theories.
3.3.1.3. Amount of Reuse Metrics
Systematic review results shows that approximately 40% of the studies represented their metrics/models/methods through the mathematical means and about 60% of the studies represented through diagram or tables or theories.
3.3.1.4. Failure Modes Models
Only two studies were found in this category and those two studies represented their models using diagrammatic, tabular and theoretical approaches.
3.3.1.5. Reusability Assessment
Systematic review results shows that approximately 28.6% of the studies represented their metrics/models/methods through mathematical means and about 71.4% of the studies represented through diagram or tables or theories.
3.3.1.6. Reuse Library Metrics
Systematic review results shows that approximately 14.3% of the studies represented their metrics/models/methods through mathematical means and about 85.7% of the studies represented through diagram or tables or theories.

4. CONCLUSION:
Assessing the value of reuse is a major concern in the software industry. For assessing its value, reuse should be measured by using the metrics and models. Measuring reuse will help the organization to know their progress in software reuse, to know how much amount of reuse is done or to assess the cost benefits of software reuse etc. Our observations shows that, the reuse metrics and models are divided in to six categories, based on their application to different areas of software reuse. The organizations can use these metrics and models for measuring reuse and reusability. As shown in the analysis section 3, the percentage of validated studies is less than the percentage of non validated studies. Out of 50 studies, 18 studies (36%) are validated and 25 studies (50%) are non-validated. Out of these validated studies, 39% are industrially validated, 56% are academically validated and 5% are validated through surveys. Among the found six categories, cost benefit analysis, maturity assessment, amount of reuse metric areas are more focused or concentrated more than the other categories. We have also tried to add some more subcategories in the taxonomy we presented, to the already existing 6 categories of reuse metrics and models. A good research is going on in this field, but it is a not sufficient.

5. ACKNOWLEDGEMENT
We thank our Associate Professor Dr. Mikael Svahnberg, School of computing, BTH, Sweden for his continuous support, and valuable inputs.

6. REFERENCES


