“Customer Satisfaction Metrics and Models”

Sunita Chulani
Sunita@us.ibm.com
IBM Research

Following is a paper that was presented at the 12th Annual ESCOM (European Software Control and Metrics) conference held in April 2001 in London, England.
Deriving a Software Quality View from Customer Satisfaction and Service Data

Sunita Chulani (sunita@us.ibm.com) & P. Santhanam
IBM T.J. Watson Research Center

Darrell Moore, Bob Leszkowicz, Gary Davidson
IBM Software Group

Abstract
Most quality and cost models use defect density to represent software quality. Customer's quality expectations are not typically based on size and complexity of the product they buy and their satisfaction can be influenced substantially by other product attributes that are not typically mapped to defects (e.g. Ease of installation and use, timely support, etc.). Consequently, new ways to measure customer view of quality are needed. In this paper, we provide analyses of customer service and survey data from eight products and discuss some key insights to lay the foundation for a better understanding of customer view of software quality. We believe that this approach\(^1\) can help us identify actions in software development and support that will address the concerns of our customers and improve their satisfaction with our products.

Keywords: customer satisfaction, development metrics, service metrics, marketing metrics, in-process, software quality

1. Introduction
Software economics and measurement research work in recent years has focused on building models for software cost, schedule and quality prediction (for example, USC's COCOMO II suite \(^1\), Putnam's SLIM \(^2\), Jones' KnowledgePlan \(^3\), etc.). The extent of the quality prediction is in determining the defect density of the product when its fully operational in the field in terms of defects/LOC (Thousands of Lines of Code) or defects/FP (Function Point), where a defect is viewed as problem resulting in some change of code distributed to the customer. While this metric may represent quality adequately from the development point of view, a typical ‘customer’ for software does not take into account the size or complexity of the product in judging the quality of the product. In addition, defect density does not really capture the overall customer satisfaction experienced by the users of the software product. For example, when a customer has difficulty with product installation, he calls in for service and this user problem is not necessarily coded as a defect as it may not need a change in the code base. Thus, even though the customer may not be satisfied, the currently available quality/cost models don’t incorporate this dissatisfaction in any way.

\(^1\) Note: This paper does not capture all the processes/analyses done in a large organization. The field of customer satisfaction and software quality is a broad and complex field and this paper makes an attempt to uncover some useful insights that may help us manage our products in a more beneficial way.
There has been some prior work (e.g. [4]) establishing some link between service and customer satisfaction metrics. Yet, more effort is needed to define meaningful and practical ways to measure the impact of the entire software development life-cycle, including the service phase, on overall customer satisfaction when the product is fully-operational in the field. This paper attempts at building this bridge between development, service and in-field customer satisfaction. It presents analyses that can be used to help determine quantitative answers (not just inclinations or expert opinion) to questions such as:

(i) What are the top 3 attributes customers care about the most?
(ii) What is the relationship between kinds of customer problems (e.g. Duplicate problem report, problem affecting a data record, etc.) and product attributes, i.e. Performance, Reliability, Usability, etc.?
(iii) Does the defect density or the time to resolution have a higher impact on overall customer satisfaction?
(iv) Does customer satisfaction improve if time to resolve severity 1 field defects is reduced?

Section 2 provides some background for the reader; section 3 details some analysis examples that help the reader understand techniques that help answer questions of the sort raised above. Section 4 provides the conclusions and future research.

2. Background

Over the past year, we have commissioned a study with participants from different groups within our organization encompassing most of the entities of the software product life-cycle, in particular, product development teams, service teams and customer satisfaction analysts. The product development teams are the primary producers of the product; the service teams work with the customers once the product is generally available and helps the customer if the customer experiences any problems with the product; the customer satisfaction analysts capture and analyze several different marketing aspects of the organization and products such as brand name satisfaction, trends in the marketplace with respect to product customer satisfaction and its relationship with competitor products’ customer satisfaction, etc. Often times, these different groups are separate entities with different focus areas. The theme of this study is to recognize and exploit the benefits of using data and information across these teams together to make better decisions to improve customer satisfaction. The reader should note that this paper presents a subset of activities/analyses done in a large organization.

Customer satisfaction that is addressed in this paper is related to product attributes that manifest as defects in software. Other aspects of customer satisfaction, such as functionality provided vs. customer expectations, price considerations, vendor relationships, etc. are not directly addressed in this study. Thus it is really from the point of view of the IT professional in the customer organization developing and supporting applications or servers using IBM software. They do not typically make the decision to buy the software and they are also not the end user of the software that is being deployed.
2.1. Data sources

Quality is viewed from three chronological views as shown in Figure 1 [see 5 for an overview of quality views, metrics and models]. Two of the views are formed by the data that were available for this study i.e. (i) data from field defects from our service teams and (ii) customer satisfaction data from our marketing teams.

![Figure 1: Quality Views](image)

When the product is being developed, several defect detection/prevention/removal activities are employed which detect in-process defects [6]. Once the product is ready and shipped, customers encounter problems and call into the product service centers for help. These problems that cause a change in the product are recorded as field defects by the service personnel. When the product is in the field, marketing teams do surveys to capture overall customer satisfaction of the product. The timeline of when what data is collected is shown in Figure 2. All these different pieces of data are used to monitor progress and identify areas of improvement.

![Figure 2: Timeline of quality-related measures](image)

The next few paragraphs describe in detail the two sources of data available for this study. It is also important to note that the processes and the associated data collection do not necessarily apply to all of IBM's business units thus limiting the scope of this work.

2.1.1. Service data
Once an IBM product is generally available, customers who have problems with the product contact our call management center and are connected to the appropriate technical support staff that helps them resolve the issue either in a preventive or a corrective way. During this process, a lot of data gets recorded in IBM’s defect databases. This data is used as a means of communication between the teams and also for analysis to improve the overall service process.

For example, when a customer calls in with a problem, the service representative opens a 'Problem Management Report' or more commonly known as a PMR at IBM. The service representative then works with the customer and/or the development team to resolve the problem. The resolution of the problem may or may not cause a change in the product. If a change is made, the problem is termed as a 'Defect Oriented Problem' and is coded as an 'Authorized Program Analysis Report' or an APAR. The change is a PTF or a 'Program temporary Fix' until it is incorporated in the next release. If a change is not made, the problem is a 'Non-Defect Oriented Problem'. The process is illustrated in figure 3. For this study, we had service data for 8 products.

![Service Process - High Level View](image)

### 2.1.2. Customer Satisfaction Survey Data

Customer satisfaction is measured by marketing groups that survey the customer base and capture overall customer satisfaction in several dimensions. The key dimensions used at IBM are popularly known as CUPRIMDSO (C: Capability, U: Ease of Use, P: Performance, R: Reliability, I: Ease of Installation, M: Maintainability, D: Documentation, S: Service/Support, O: Overall Satisfaction). In addition to these, other dimensions such as Price, Availability, etc. are also used.

For each of the dimensions, a qualitative scale is often times used which has ratings such as (i) Very Dissatisfied (ii) Dissatisfied (iii) Neutral (iv) Satisfied (v) Very Satisfied. Net Satisfaction Index is computed using the responses on the rating scale as a weighted satisfaction index.
For the purposes of this study, we had customer satisfaction data for 8 products. A lot of marketing research has been done to study customer satisfaction data [7, 8, 9]. But most of these efforts are limited just to the marketing data source and do not establish links between other quality views to exploit all the information to complete the picture.

3. Analysis Results

As described above, for the purposes of this study, we had marketing data for 8 products that are very often used together. Detailed analysis on the data gave some interesting insights in the customer satisfaction of this set of products. Some of the highlights of the analysis are presented in the next few paragraphs.

3.1 What are the top 3 attributes customers care about the most? What are we doing to address these attributes?

One question on the customer satisfaction surveys that we ask is: “What quality characteristics come to mind when considering a high quality software product on this operating system?” Answers to this question are in the form of free text and customers can say as much as they wish. Sometimes they have a long list of characteristics and sometimes just a single attribute. This data gives us the most direct response from our customers to improve customer satisfaction. Using our data and text mining on the free text answers we can understand what quality characteristics are important to our customer and what we should focus our efforts on. Such analysis may reveal that in addition to CUPRIMDS, other attributes such as Speed, Compatibility, Price, etc., may also be relevant.

Using this knowledge that we can extract from our data and by doing some additional mining using our marketing data, we can further analyze if we are indeed satisfying our customers in terms of the top attributes that are revealed as relevant when analyzing this data. For example, for each of the CUPRIMDS attributes we ask our customers two questions: (i) How satisfied are you with the attribute for this product and (ii) how important is this attribute. This data can help us do gap analysis to understand if we are meeting our customers expectations.

For the set of 8 products, let’s say attribute A is the most relevant characteristic our customers care about. Gap analysis on attribute A can help us answer the question: “Are we meeting our customer’s expectation for attribute A?” Figure 4 illustrates the Satisfaction vs Importance ratings for attribute A.
From the above chart, we can clearly see that for attribute A, our data distribution indicates that our customers are not getting the same level of Satisfaction as they think is Important. Hence, we really need to focus our efforts, in reducing the height of the SAT<IMP bar on attribute A and other high impact attributes. What actions can we take earlier in the life of the product? For example, using our service data to improve overall customer satisfaction? We will discuss this in the next section.

On the other hand, consider a different scenario: say the gap analysis results showed the distribution depicted in figure 5. In such a scenario, we are meeting our customer’s expectations as the frequency of SAT<IMP is much lower than that of SAT>IMP.
3.2. What is the Relationship between Customer Problems and Product Attributes?

Let's look at an example and see how we can use our Marketing data and Service data to improve the Satisfaction on some product characteristic, say Reliability. This example can be used to do analysis to answer the question: What is the relationship between kinds of customer problems (e.g. Duplicate problem report, problem affecting a data record, etc.) and product attributes i.e. Performance, Reliability, Usability, etc.?

As explained in section 2, when a customer has a problem they contact our service teams. The service team help resolve our customers problems and in the process they record a lot of detailed information about the problems encountered in-field. If we can analyze the defects to understand which types are highly correlated to Reliability and resolve these defects more efficiently then we have a chance of improving the Satisfaction on Reliability.

Doing a correlation study on our data, we find that Reliability is most highly correlated to defects which are classified as “Records closed as they are duplicates of open records”. Similarly, Reliability is less correlated to records which are classified as “Change team could either not reproduce the problems and the problems had already been resolved”. Based on this information, we know that as soon as we have a customer-reported defect that is a duplicate, we focus our efforts on resolving it soon, if we want to focus our efforts in improving satisfaction with Reliability. Similarly, our data shows that our customers are usually quite satisfied in terms of Reliability when they call in with a problem and our teams respond with a fix (the problem has been fixed). Hence, this process helps us focus our service efforts in the right direction to improve a customer satisfaction attribute (such as Reliability) based on actual information that is gathered by our service teams.

3.3. Does defect density or time to resolution have an impact on Customer Satisfaction? Does Customer Satisfaction improve if time to resolve severity 1 field defects is reduced?

The next few paragraphs present ways for the reader to look at different data that is available in any large organization to facilitate better decision-making. Let's try to answer the question: Does the defect density have an impact on overall customer satisfaction?

Table 1 shows the Overall Satisfaction and # of APARs for the 8 products for 2 years. The table is sorted in the order of satisfaction index of 'O' i.e. Overall Satisfaction with the product. This table shows no relationship between the # of APARs and O. Common sense would say that as the # of APARs increases, Overall Satisfaction should decrease but this is clearly not true for all the products. Further analysis of this trend shows that some of the products have been in the market longer than others and the extent of usage explains the higher number of APARs in some products that have high Overall Satisfaction (for example, product G). The metrics “# of APARs/Thousand
Lines of Code" or "# of APARs/Function Point" may be better to use for such an analysis but this information wasn’t available for all the products in this sample set.

Table 1: Trend of Overall Satisfaction and Total # of APARs*

<table>
<thead>
<tr>
<th>Product</th>
<th># O</th>
<th>Total # of APARs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prod A</td>
<td>69</td>
<td>17</td>
</tr>
<tr>
<td>Prod B</td>
<td>69.4</td>
<td>15</td>
</tr>
<tr>
<td>Prod C</td>
<td>70.9</td>
<td>167</td>
</tr>
<tr>
<td>Prod D</td>
<td>75</td>
<td>36</td>
</tr>
<tr>
<td>Prod E</td>
<td>75.6</td>
<td>455</td>
</tr>
<tr>
<td>Prod F</td>
<td>77.5</td>
<td>100</td>
</tr>
<tr>
<td>Prod G</td>
<td>80.8</td>
<td>5125</td>
</tr>
<tr>
<td>Prod H</td>
<td>82.6</td>
<td>278</td>
</tr>
</tbody>
</table>

Now, let's determine if the time to resolution of defects has an impact on satisfaction? Let's consider 3 of the 8 products which have steadily increasing Overall Satisfaction ratings. For example, let's consider products D (O = 75), F (O = 77.5) and G (O = 80.8). Product F’s O is about 3.3% higher than Product D’s O and Product G’s O is about 4% higher than Product F’s O. Now, let's see if the time to resolution of a subset of field-reported defects, say Severity 1 APARs, has any impact on the steadily increasing O for these three products. Table 2 illustrates the # of Severity 1 APARs, the percentage of Severity 1 APARs, average time to resolve a Severity 1 APAR and Overall Satisfaction for products D, F and G.

Table 2: Trend of APAR severity, density, time to resolution*  

<table>
<thead>
<tr>
<th>Product</th>
<th># of Sev 1 APARs</th>
<th>% of Sev 1 APARs</th>
<th>Average Time to Resolution</th>
<th>Overall Satisfaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prod D</td>
<td>3</td>
<td>8</td>
<td>129</td>
<td>75</td>
</tr>
<tr>
<td>Prod F</td>
<td>15</td>
<td>15</td>
<td>111</td>
<td>77.5</td>
</tr>
<tr>
<td>Prod G</td>
<td>943</td>
<td>18</td>
<td>69</td>
<td>80.8</td>
</tr>
</tbody>
</table>

* Please note that the defect data in tables 1 and 2 has the trends of the original data but the numbers have been multiplied by a constant to assure confidentiality.

Table 2 clearly shows that even though the number (and percentage) of Severity 1 APARs increases, the decreasing time to resolution of these APARs results in the Overall Satisfaction to steadily increase. This helps our teams focus efforts on reducing the time to resolution of defects if our goal is to increase Overall Satisfaction.

Several other analyses have been done using the service/marketing data to improve quality and customer satisfaction of our products in ways similar to the ones shown in the above paragraphs. Such analysis helps us answer the questions posed in the abstract. The aim of this paper was to show some of the ways analysis can be done to improve our products. We strongly believe that analyzing historic data is a very good way to better manage our products.

4. Conclusions and Future Research
This paper presented examples of analysis that illustrate the relationship between development and service metrics (number of defects discovered, time to resolve a defect, severity of defects, etc.) and customer satisfaction survey metrics (overall customer satisfaction, top attributes customers look for in a product, etc.). Data from 8 different products is analyzed using several analytical techniques in a systematic way to help answer many important questions related to improving overall customer satisfaction.

A lot of research needs to be done in this broad and complex area of understanding our customers and their level of satisfaction with our products. We are continuing to do more work in this area to further improve our product quality and customer satisfaction. Some of the areas that we are actively working on include: understanding satisfaction with the service process once the customer problem has been resolved, determining a solution-based versus product-based view of customer satisfaction, etc.

5. References


