CHAPTER 1

INTRODUCTION AND DESIGN OF THE STUDY

1.1 INTRODUCTION

Open Source Software (OSS) has made a major impact on the software industry, all over the world, both developed and as well as developing countries. Basically, the Open Source Software Development (OSSD) is virtually a distributed. The method of development with the involvement of volunteers as software developers, who are located geographically in various regions. Since the decade of 1990s, (Michlmayr 2005) the Open Source Software Project Management (OSSPM) is facing the challenges of globalization in terms of defining the software development process and transforming local phenomena into global phenomena in both development and deployment process. Besides a steep increase in the number of volunteers being actively involved in open source software development, the OSS concept offers not only low cost of technology acquisition, but also an efficient scheme of cooperation to exploit such technology.

In a broader perspective, open source software is developed in a collaborative manner and most of the developers are volunteers. Although this collaborative method of development has created a substantial amount of software, the software development procedure is often described as unstructured and unorganized.

The most well-known and successful open source projects such as ‘Apache Web Server’ and ‘Linux Operating System’ have confirmed the attractiveness of open source development as a new method of producing
software. While open source software has proved to be popular and is often found to be better quality than proprietary software, it is still not clear as to what factors contribute, either directly or indirectly to this quality. Some existing works investigate OSS phenomenon in a quality improvement perspective and examines where enhancements are possible to the development process.

This form of volunteer cooperation in developing OSS technologies and in customizing such technologies can aid the developing countries in improving their current technology transfer systems. This is evidenced in the evolutionary process of OSSPM itself.

1.2 NEED FOR THE STUDY

While the conventional Closed Source Software Development (CSSD) is a proprietary based approach, OSSD is a volunteers community based approach for respective common entities. Given this aspect of difference, the success measures engaged in CSSD may not essentially fit into OSSD environment (Crowston and Scozzi 2002; Stewart et. al. 2006). To further explore this OSSD in context of release process, the study classifies open source software projects into two types based on the application usage, viz. Developer oriented projects and User oriented projects. The developer oriented projects aim at development of a software component that can be customized and re-used by other developers. These projects enable developers to extend the existing software by extracting the works of the previous developers through recycling the source code.

The user oriented projects aim for developing a system that primarily provides utilities to end-users in various fields of applications. Their main objective is to provide end-users with functionalities that are either difficult or expensive to obtain through proprietary software.
Thus, from the above explanation, it can be inferred that basic differences do exist between developer oriented projects and user oriented projects. Project type plays a moderating role in the relationship between project management and project performance. In other words, to make an impact on the performance of the OSS project the project management has to be aligned with project type. Project managers, therefore, need to place more emphasis on various factors that influence the release approach based on the intended usage of projects and uniqueness of the project community.

OSSPM consists of technical as well as managerial activities which are needed to deliver software products. These are the issues related to process management, release management and coordination management. This area of study that received little attention in OSS research so far. In OSSD, release management becomes even more significant and difficult than in proprietary software development. The OSS community is dynamically changing, as developers and users often join and leave projects constantly. An effective release management approach can help in involvement and retention of members in OSSD. In other words, a release is a growth phase of a software application, evolved during the software development process. Release management is an important part of quality management in OSSD, since it is concerned with the delivery of high quality software to developers and end-users. Previous studies have investigated the time based release approach which is a feasible alternative to feature-driven approach. But, time based release is not the only approach to identify the quality of an OSS.

Therefore, this study investigates the factors that are linked in the release approach as a measure for the effective implementation of the release management in OSSD.

1.3 OBJECTIVES OF THE STUDY

- To study the open source software development in general and the product evolution and product quality in particular.
To analyze the release process in ‘Closed Source Software Development’ and ‘open source software development’.

To assess the impact of release management in open source software development.

To ascertain the factors that influence the release management.

To identify the factor based release approach which influences the quality improvement of OSSD.

To suggest a model for release approach in open source software project management.

1.4 STATEMENT OF THE PROBLEM

The open source Software is developed to retrieve the outcome and release with high quality software as per the expectation of the computer users. The expectations of the computer users have no limit and their needs are fulfilled by developing various novel softwares. But in reality, it is highly difficult to ascertain the exact needs and expectation, because when they fulfill a particular software need, automatically a new need emerges with high expectation in usage of more user-friendly software. Having observed the various dimensions of needs of the computer users, a new model was developed by exploring the impact of release management through world class quality.

The expectations and goals of the computer users are totally uncertain and hence, to fulfill the emerging needs of the users, a new model has been developed with various domain of knowledge management.

The Project Management Institute, USA has identified nine areas of knowledge that a project manager needs. These include integration management, scope management, time management, cost management,
quality management, human resource management, communication management, risk management, and procurement management (Duncan 1996).

The emerging new goals among the computer users are usually provided in the form of a release process, where the development of open source software model is an incremental process moving towards a series of sequenced release, covering high quality consciousness in open source software development. It is an integrated approach model covering four major factors of knowledge management domain viz, Human resource management, time management, communication management and quality management.

So for the open source software developers have concentrated only on time based and feature based approach. It totally lacks in concentration towards quality and human resource. This was identified as a major gap in this research and in order to fulfill the gap, this integrated approach covering four major domains are developed for open source software users, who in turn may become a open source software developer. The present software development highly influences the release management process in OSSD which has been considered one of the main areas of the release approaches in terms of the quality improvement in open source software project management.

The impact of release management on quality improvement in OSSPM is an area which is largely unexplored. The OSS development processes are influenced by new factors of release approach. To ensure quality improvement, there is a need to evolve a different approach to overcome the deficiencies of the existing release approaches. The aim of the research is to ascertain relevant information from OSS volunteer’s community on the influential factors towards the release management activities.

To bridge the geographical distance of the wide spread volunteers, a different release approach involving the characteristics of quality and its
implications in OSSPM necessitates empirical investigation. The influence of certain factors like electronic mode of communication has the necessary potential to bridge the distance among wide spread volunteers community. In the release management activities, it has implications for quality improvement in OSSPM. The probing of this dimension is justified in terms of the formulated research questions and objectives of the present study.

The application mechanisms in release management activities and its derivations affect the quality of OSSPM. Analysis of this aspect serves as the process to evolve a new approach for quality improvement of OSS for the benefit of the volunteer’s community.

1.5 RESEARCH QUESTIONS

The impact of release management is not only on the quality, but also on supervising how changes flow in the software system with each phase of the subsystem (Human Resource, Communication, Quality and Time Management). The present study of release management generally highlights the management efforts which includes a great deal of coordination of the selected subsystems and solves the most complex task at user friendly level.

In the context of the above issues, the following research questions are addressed.

1) What are the major factors involved in release approach?

2) To what extent the selected subsystems have influenced on quality improvement in open source software project management?

3) How for the integrated approach is feasible to solve the most complex task?
1.6 SCOPE OF THE STUDY

Factors related to release management in OSS projects and its impact on quality improvement in OSSPM will be the main focus of this research. The present study aims at analyzing release management activities and factors that influence release management towards quality improvement in order to evolve a model for release approach in OSSPM. This study also explores whether OSS developers (consisting primarily of volunteers) collaborate through the Internet to produce not only software of high quality, but also sustain their development.

The framework of the study includes the core issues centering around the release management in OSSD. Dependent variable viz., the level of quality maintained in OSSD by the volunteer’s community with the related independent variables. Similarly, an attempt was made to measure the task achieved by the developers in the context of respective dependent variable. The approach is specifically confined to ascertain the impact of release management on quality improvement.

The study will endeavor to concentrate on OSS developers located in the southern region of INDIA, viz. Chennai and Coimbatore in Tamil Nadu State. The State of Tamil Nadu is purposively selected due to its extensive involvement in developing software companies and expanding their business network. In Tamil Nadu, Chennai and Coimbatore are the two cities which mainly concentrate in the rapid development of IT sector. This empirical research study for data gathering and analysis covers 400 respondents, i.e. 200 respondents (OSS Developers) representing Chennai and 200 respondents (OSS Developers) representing Coimbatore.
1.7 RESEARCH METHOD-TOOLS AND PRACTICES

1.7.1 Sampling

For the present study, Tamil Nadu State was intentionally selected due to their aggressive involvement in developing the software companies and its business network throughout the globe. There are two cities in Tamil Nadu viz. Chennai and Coimbatore which mainly concentrates in the rapid development of IT sector. For the purpose of this empirical research study, the respondents were chosen from these areas. The names of the software companies using open source software were identified from the web resources and the sample respondents were chosen randomly from these IT sectors. Since, Chennai and Coimbatore cover a thick population with substantiate number of IT companies, adequate number of sample respondents were selected to ascertain their views on open source software project management. The researcher has circulated the finalized questionnaire to 500 respondents. Among these 500 respondents 430 volunteers have responded properly. The researcher has crosschecked all the data and again eliminated the unfilled information as well as bogus information. Finally it has been restricted with 200 samples from each city. The geographical distribution of the respondents using open source software is given in the following Table 1.1.

Table 1.1 Geographical representation of sample respondents

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Name of the City</th>
<th>No. of Respondents</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>1</td>
<td>Chennai</td>
<td>97</td>
<td>103</td>
</tr>
<tr>
<td>2</td>
<td>Coimbatore</td>
<td>30</td>
<td>170</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>127</td>
<td>273</td>
</tr>
</tbody>
</table>
1.7.2 Data Collection

The validity of any research is based on the systematic method of data collection and analysis. In the present study both primary and secondary data were used.

1.7.2.1 Primary data

Primary data were collected from the respondents who were using open source software in the study area. First-hand information was collected from four hundred sample respondents living in Chennai and Coimbatore. The data were collected with the help of a well-structured questionnaire from the selected sample respondents. A questionnaire contains various information like, socio-economic data of the respondents, awareness of open source software, extent of utilizing the software, level of satisfaction, effective communication, quality of open source software and the opinion on open source software facilities and services and various other kinds of problems. These were organized in the systematic way. The data thus collected from the open source software were organized in a simple tabular form.

1.7.2.2 Secondary data

The primary data were supplemented by a spate of secondary sources of information. In order to learn about the open source software concept, several literature reviews were collected from well-equipped libraries in Bangalore, Chennai, and PSG Learning Resource Centre, Coimbatore. A number of standard text books and journals were studied to obtain pertinent literature and conceptual framework of open source software. Further, internet
web resources were also used to collect the latest information of organized open source software usage.

1.7.3  Pre-test

The questionnaire meant for the respondents was pre-tested with forty respondents who were using open source software. After pre-testing, a few respondents were suggested a few changes in part-IV in the questionnaire release management activities: Testing Process and Defect Handling process. These necessary modifications were made in the questionnaire to fit in the track of the present study.

1.7.4  Frame Work of Analysis

The core of the study being “The release management in open source software development”, the study centers around the dependent variable viz., the level of quality maintained by the users of open source software with the related independent variables. Similarly, the tasks achieved by the users were analyzed with the respective dependent variables.

1.7.4.1  Approach to the extent of utilizing the open source software

The difference in the extent of utilizing the open source software and their level of quality maintained among the users were studied based on their Age, Gender, Education, Occupation, Income Level, Family Size, Type of Family and Wealth Positions of the respondents. These were studied by means of Two-way tables, Percentages, Averages, Ranges and Standard Deviation.

1.7.4.2  Chi-square test

The degree of influence of the following independent variables pertaining to respondents’ level of quality maintained by the open source
software developer, such as (i) Respondents’ Age, (ii) Respondents’ Gender, (iii) Respondents’ Education, (iv) Respondents’ Occupation, (v) Respondents’ Income, (vi) Respondents’ Family size, (vii) Respondents’ Wealth position, (viii) Respondents’ Type of family, (ix) Respondents’ Level of Awareness and (x) Respondents’ Opinion on open source software were studied using chi-square test.

In order to identify the factors influencing the level of quality maintained by the open source software developers in the selected software companies in the study area, a Chi-square ($\chi^2$) test was applied and the formula is given below.

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

With Degree of Freedom (D.F.) = (c-1) (r-1) where, O = Observed frequency, E = Expected frequency, c = Number of Columns (Dependent variable classification) and r = Number of Rows (Independent variable classification).

1.7.4.3 Multiple regression analysis

Regression is a statistical relationship between two or more variables. When there are two or more independent variables, the analysis that describes such relationship is multiple regression. This analysis is adopted where there is one dependent variable that is presumed to be a function of two or more independent variables. In multiple regression, a linear composite of explanatory variables is formed, in such a way that it has maximum correlation with an active criterion variable. The main objective of using this technique is to predict the variability of the dependent variable, based on its co-variance with all the independent variables. It is useful to predict the level
of dependent phenomenon through Multiple Regression Analysis models, if the levels of independent variables were given. Since the dependent variable, the “level of quality” is qualitative in nature; it is measured by selecting ten statements and Likert’s five point scaling system. Scale scoring technique was employed to convert the quality metrics in to quantification. The linear multiple regression problem is to estimate coefficients of $\beta_1, \beta_2, \ldots, \beta_j$ and $\beta_0$ such that the expression,

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_j X_j$$

provides a good estimate of an individual $Y$ score based on the $X$ scores,

where,

$Y =$ Level of quality maintained by the respondents  
$X_1 =$ Respondents’ Age  
$X_2 =$ Respondents’ Gender  
$X_3 =$ Respondents’ Education  
$X_4 =$ Respondents’ Occupation  
$X_5 =$ Respondents’ Income  
$X_6 =$ Respondents’ Family Size  
$X_7 =$ Respondents’ Wealth  
$X_8 =$ Respondents’ Type of family  
$X_9 =$ Respondents’ level of Awareness on Open Source Software  
$X_{10} =$ Respondents’ Opinion on Open Source Software

and $\beta_1, \beta_2, \ldots, \beta_j$ are the parameters to be estimated.
1.7.4.4  Garrett ranking technique

This technique was used to rank the problems faced by the respondents using open source software in the study area. In this method, the respondents were asked to rank the given problems according to the magnitude of the problem. The order of merit given by the respondents was converted into ranks by using the following formula.

\[
\text{Percentage Position} = 100 \left( \frac{R_{ij} - 0.5}{N_j} \right)
\]

\( R_{ij} = \) Rank given to \( i^{th} \) item by \( j^{th} \) individual

\( N_j = \) Total number of population

The percentage position of each rank thus obtained is converted into scores by referring to the table given by Henry Garrett. Then for each factor, the scores of individual respondents were added together and divided by the total number of respondents for whom the scores were added. These mean scores for all the factors were arranged in order of ranks and the inference were drawn.

1.7.4.5  Factor analysis

Factor Analysis was used to study a complex product or service, in order to identify the major characteristics or factors considered important by the respondent. The purpose of factor analysis is to determine the responses to the several numbers of statements, which are significantly correlated.

1.8  OPERATIONAL DEFINITION

The following operational definitions are presented in an orderly form
a) Source Code: Source code implies a set of instructions written in a computer language in the form of computer software for specific application usage.

b) Open Source Software (OSS): It is “a type of software whose source code is released by the impact of multiple factors and to be made available to volunteers/ computer users to enable them to modify with focus on quality and redistribute without any constraints”

c) Open Source Software Development (OSSD): The OSSD is an incremental method, in which the efforts of volunteer community are involved virtually to develop a source code in order to evolve the projects.

d) Open Source Software Project Management (OSSPM): It is related to managing multiple factors in OSSD which includes Human resource, Communication, Time and quality.

e) Volunteer Community: Volunteer community is groups of computer users who uses or develop software either directly or indirectly for evolving Open Source Software.

f) Release Management (RM): It is a set of activities which are related to the release of a source code of a project to volunteer community. It includes testing, debugging, verifying quality metrics of a project.

g) Release: A release is a gathering of new and/or changed set of instructions that are linked to form a new versions of an evolving project.
h) Quality Improvement: It is an improvement process of a OSS project related to functionality, reliability and efficiency of a OSS project.

i) Human Resource: Human resource refers to the volunteer communities who are geographically wide spread and are involved in OSS project activities.

j) Communication: It is an information exchange process through electronic mode of communication for effective collaboration among volunteers.

k) Release Schedule: It means a time schedule with deadlines for evolving a project.

l) Feedback: It is a process of obtaining information on OSS release management activities by OSS developers.

m) Software evolution: The tendency of software to change/grow overtime.

n) Committer: A committer is also a developer who has been recognized by the community and given write access.

o) Developer: A user who contributes to the project in the form of code or documentation.

p) User: Someone that makes use of software also called as an end-user.

1.9 LIMITATIONS OF THE STUDY

The sample is confined to only two cities located in Tamil Nadu State in India. Hence, the results arrived from the study may or may not be applied to other States. Further, survey method was adopted to collect data for
the study, which may have its own limitations. However, in order to make the results reliable for drawing conclusions relevant for the universe of the study, care had been taken to minimize the bias, through cross checks, whenever the accuracy and reliability of the data given by the respondents were doubted. In addition to release management, the exploratory study has revealed other two areas of interest namely Process Management and Coordination management with regard to quality management of open source software project management. These two areas require further investigation, covering more number of States.

While the above paragraphs reflect an introductory framework of the research, the next chapter presents a review of relevant literature on OSS, OSSD, OSSPM and quality process in the context of release management activities.

1.10 STRUCTURE OF THE THESIS

The present study is organized into five chapters:

- The first chapter presents a clear picture of research design. it includes introduction, need for the study, statement of the problem, objectives of the study, research methodology adopted, framework of analysis, scope of the study, period covered by the study and limitations of the study.

- The second chapter presents the review of relevant literature.

- The third chapter presents a theoretical framework of software development process.

- The fourth chapter covers the integrated approach of release management in open source software project management.

- The fifth chapter presents the summary of findings, suggestions and conclusion.