TASK ALLOCATION IN GLOBAL SOFTWARE DEVELOPMENT: AN EMPIRICAL STUDY

BY

SAJID ANWER

A Thesis Presented to the
DEANSHIP OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

COMPUTER SCIENCE

May 2015

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN- 31261, SAUDI ARABIA

DEANSHIP OF GRADUATE STUDIES

This thesis, written by **Sajid Anwer** under the direction his thesis advisor and approved by his thesis committee, has been presented and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of **MASTER OF SCIENCE IN COMPUTER SCIENCE.**

Thesis Committee

Thesis Advisor

Dr. Sajjad Mahmood

Member

Dr. Mahmood Khan Niazi

Member

Dr. Mohammad Rabah Alshayeb

Department Chairman

Dr. Abdulaziz Alkhoraidly

Dean of Graduate Studies

Dr. Salam A. Zummo

Date 3/6/15

© Sajid Anwer

2015

IN THE NAME OF ALLAH, THE MOST BENEFICIENT THE MOST MERCIFUL

Dedicated To

My respected parents, brothers, teachers and friends whose care, love and guidance have been enlightening my life since my first breath.

ACKNOWLEDGEMENT

All gratefulness and praises are to Almighty Allah who bestowed me the illumination and courage to fulfill my commitment towards my research thesis. Peace and blessings upon the final Prophet Muhammad (Peace be upon him), his family, and his companions.

I want to thank KFUPM for giving me the opportunity to pursue graduate studies and providing tremendous research facilities and financial assistance during the course of my MS program.

It is matter of great pleasure and honor to express my deep sense of gratitude for the continuous guidance and precious advices extended by my project supervisor Dr. Sajjad Mahmood. He has been a constant source of enormous motivation and enthusiastic support throughout the span of the thesis. Due to his compassionate benefaction and competent coaching, I have been able to finish this thesis.

I am also thankful to Dr. Mahmood Niazi and Dr. Muhammad Alshayeb for their support, productive criticism and comments which were of utmost significance. Their valuable suggestions and friendly counseling backed me in achieving my objectives.

I then would like to extend my deepest gratitude and respect to my parents, my siblings, respected teachers Mr. Muhammad Waheed Aslam, Dr. Moataz Ahmed and friends Yasir Abdullah Siddiqui, Muhammad Imran Qureshi, Waleed Umar, Rashid Maqsood and Samran Naveed who deserve my deepest regards for supporting me whenever I needed them always been a source of encouragement, motivation and love.

Table of Contents

LIST OF TABELS

LIST OF FIGURES

THESIS ABSTRACT (ENGLISH)

Name	Sajid Anwer
Title	Task Allocation in Global Software Development: An Empirical Study
Degree	Master in Computer Science
Major Field	Computer Science
Date of Degree	May, 2015

Over the last decade, many firms in the world have started global software development (GSD). In GSD, a company (client) contracts out all or part of its software development activities to another company (vendor), who provides services for remuneration. Previous work suggests that half of the companies that have tried GSD have failed to realize the anticipated outcomes which have resulted in poor outsourcing relationships, misunderstanding of projects" requirements, high costs and poor services. In order to address the risk and problems associated with GSD, determinants from industry experience need special consideration to address unique characteristics of geographically distributed software development model. One critical factor of GSD for successful project completion with in specified time and budget, is the allocation of tasks as project managers not only need to consider her/his workforce but also need to take into the account the characteristics of the sites, their relationships and task characteristics.

The objective of this research is to identify factors, from the client and vendor perspective, that influence task allocation in Global Software Development Projects. We follow three phase approach, we first identify the factors through carrying out Systematic literature

review. In the second phase the identified factors are validated using an industrial empirical study. We identified factors such as site technical expertise, time zone difference, resource cost, task dependency, vendor reliability, task size and vendor maturity level as key task allocation factors in globally distributed software projects. We validate literature findings from industry experts' using statistical approaches (Chi-Square Test, Spearman's Rank order Correlation etc.) to analyze and compare SLR and empirical Study findings. In third phase, we develop a new task allocation model using genetic algorithm approach based on our empirical results findings. Our research findings has the potential to help software development organization in achieving the true potential of global software development.

THESIS ABSTRACT (ARABIC)

الاسم الكامل: ساجد أنور

عنوان الرسالة: توزيع المهام في تطوير البرمجيات العالمية: دراسة ميدانية

التخصص: علم الحاسب الآلي

تاريخ الدرجة العلمية: مايو 2015

على مدى العقد الماضي، قد بدأت العديد من الشركات في العالم بتطوير البرمجيات العالمية (GSD) . في GSD ، شركة (العميل) ترسل عقود كل أو جزء من أنشطة تطوير البرمجيات لشركة أخرى (بائع) ، الذي يقدم خدمات مقابل أجر. وتشير الأعمال السابقة أن نصف الشركات التي حاولت GSD فشلت في تحقيق النتائج المتوقعة التي أسفرت عن السمعة السيئة للاستعانة بالمصادر خارجية، وسوء فهم المشاريع و المتطلبات والتكاليف العالية وسوء الخدمات. من أجل معالجة المخاطر والمشاكل المرتبطة بالGSD ، عوامل الخبرة في مجال الصناعة يحتاجون إلى مراعاة خاصة لمعالجة الخصائص الفريدة للتوزيع الجغرافي لنموذج تطوير البرمجيات. أحد العوامل المهمة بال GSD لإنجاز مشروع ناجح لتسليمه في الوقت المحدد والميزانية المحددة، هو توزيع المهام كمديري المشاريع لا تحتاج فقط للنظر الى القوى العاملة فقط لكنه أيضا بحاجة إلى ان تأخذ في الاعتبار خصائص المواقع، وعلاقاتهم و خصائص العمل.

الهدف من هذا البحث هو تحديد العوامل، من منظور العميل والبائع، المهمة لتوزيع العمل في مشاريع التنمية العالمية للبرمجيات. نحن نتبع ثلاثة مراحل، أولا التعرف على العوامل من خلال تنفيذ مراجعة الأدبيات المنهجي. في المرحلة الثانية يتم التحقق من العوامل المحددة باستخدام دراسة عملية. تم تحديد عوامل مثل عامل الخبرة التقنية بالموقع ،فارق التوقيت، تكلفة الموارد و الترابط بين المهام، موثوقية بائع، وحجم المهام ومستوى نضج البائع كعوامل مهمة و رئيسية في مشاريع البرمجيات العالمي. تحققنا من صحة نتائج الأدب من استخدام النهج الإحصائية (اختبار تشي ساحة، اختبار سبيرمان الخ) لتحليل ومقارنة النتائج ال SLR مع الدراسة التجريبية. في المرحلة الثالثة، طورنا نموذج توزيع المهام الجديد باستخدام نهج الخوارزمية الجينية استنادا لنتائج الدراسة التجريبية. نتيجة بحوثنا لديها القدرة على مساعدة منظمة تطوير البرمجيات في تحقيق الإمكانات الحقيقية لتطوير البرمجيات العالمية

CHAPTER 1

INTRODUCTION

1.1. General

Global Software Development (GSD) is a modern software engineering paradigm. GSD is the process where a company (client) contracts all or part of its software development activities to another company (vendor), who provides services in return for a financial compensation [1]. Over the past 10 years, many organizations across the globe have started adopting GSD in order to reduce their software development cost. GSD helps companies to leverage the benefits of multi-site development with respect to time, cost and access to skillful resource. Software development outsourcing has been rising steadily and an 18-fold increase in the outsourcing of IT-enabled business processes is estimated[2]. Small and medium sized organizations can use outsourcing to address their issues of limited resources and lack of technical expertise. This creates a business opportunity for the Vendor organizations and hence they are struggling to contest internationally in attracting software development projects. Previous research suggests that 50 % of the companies that have tried global software development (outsourcing) have failed to realize the projected benefits due to misunderstanding the projects' requirements, poor global relationships, high costs and poor services [3, 4]. One of the major concerns is that most of the clients certify global contracts with their vendors before testing their project management capabilities readiness for the global software development activities[3, 5].

1.2. Problem Definition

There are many reasons for initiating global software development project[6, 7]. Client organizations benefit from offshore outsourcing because vendors in developing countries (offshore vendors) typically cost one-third less than onshore vendors and even less when compared with in-house operations. Amongst many other reasons for outsourcing, generally client organizations outsource their software development work to offshore locations to gain cost and quality advantages, access to leading-edge technology and the ability to focus on core competencies[7]. Moreover, offshore vendors improve their skills and service quality with the experience of offshore outsourcing projects and by learning new ways to satisfy the clients‟ needs. It is professed that offshore outsourcing vendors can add significant value to their clients‟ supply chains[8]. Conversely quite apart from the outsourcing benefits there are many risks in an outsourcing process[9], such as temporal incompatibility, cultural differences and hidden costs[10].

Cost reduction is the major motivator for software outsourcing[11, 12]. Others motivators for outsourcing include accessing leading-edge technology available at outsourcing vendor organizations and to focus on core business of the organization[12].

The results of a survey shows that eight out of every ten firms that have outsourced their software development project to an offshore vendor have faced major problems due to insufficient preparation and poor management by both the vendor organizations. Nam et al. [13] found from their investigation of 93 client companies that 36 did not plan to continue their relationships with vendors. The root cause of many failures is the increased complexity in development projects due to outsourcing. This complexity outcomes in "high coordination costs"[14], "information security problems" [15], "lack of direct communication"[16], "perceived loss of expertise in the outsourced activity"[17], "cultural misunderstandings"[18] and "infrastructure problems"[19]. Other risks are threat of opportunism, unexpected cost, trust and security concerns, geopolitical risk, and language barriers.

1.3. Research Objectives

The overarching objective of this project is to assist GSD organizations in better understanding task allocation/work distribution activity in globally distrusted development projects. This research has been designed to reduce the gap between GSD Literature and industry experts in such a way that it is accessible to both industry experts and researchers. The objectives of this research are as follows:

1. Identify the key factors which should be considered during the task allocation decisions in GSD organizations. In order to address this research objective, we have designed the following research question:

   RQ 1: What are the criteria that influence task allocation decisions in globally distributed projects?

2. Identify state-of-the-practice solutions for the identified criteria under which task allocation decisions are taken in GSD organizations? This research objective can be addressed by following research question:

RQ 2: What are the possible solutions for the identified task allocation criteria in GSD organizations?

3. Implement a task allocation model using genetic algorithms based on empirical study findings.

Addressing above objectives will assist GSD organizations in better understanding, planning and managing task allocation decisions in global software development projects. Our contribution in this project will assist other researchers with a firm basis and knowledge to develop new task allocation practices and tools that will help address the failures reported for global projects.

The above two research questions will be answered using the following research methodologies:

1. Systematic literature review (SLR).
2. Empirical study with software industry.

In order to accomplish our objectives the following tasks are performed.

1. Identify the factors that affect task allocation in GSD with the help of SLR methodology.
2. Validate our findings by doing empirical study with software industry experts.
3. Compare the findings of SLR with industrial data.
4. Implement a task allocation model in GSD using genetic algorithms based SLR and industrial data findings.

1.4. Research Approach

In order to achieve objectives we have designed an appropriate research methodology in which data will be collected from project managers as well as from the published literature (i.e. via a systematic literature review process) [20]. This two-step process will give us confidence in the reliability of the data collected. In addition to this we will implement task allocation model using genetic algorithm based on SLR and industrial data findings.

A SLR require more effort than conventional literature reviews, but provides a much stronger base for making claims to answer research questions. Technically SLR can be defined as "A systematic literature review (often referred to as a systematic review) is a means of identifying, evaluating and interpreting all available research relevant to a particular research question, or topic area, or phenomenon of interest" [21] in order to investigate specific research questions presented in section 1.3. Figure 1.1 explains and gives an overview of a Systematic Literature Review (SLR). The rationale behind doing the SLR is to identify project management challenges in GSD. We will be following the systematic literature review guidelines given by Kitchenham & charters [20].

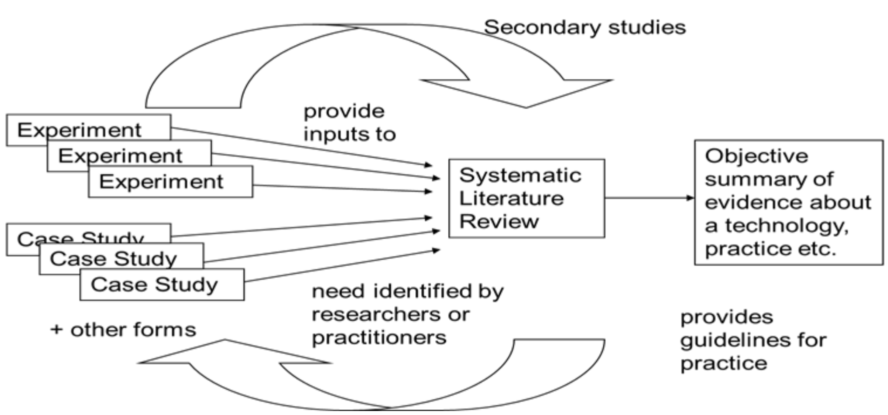

Figure 1.1 Overview of SLR

Any SLR would have the following process in a step by step manner as shown in Figure 1.2.

1. Description of a systematic review protocol
2. Defining search strategy using major terms from the breakup of the research questions.
3. Exclusion and inclusion criteria for selecting primary studies.
4. Extraction and synthesis of relevant data answering the research questions.
5. Description of quality assessment mechanisms.

Once the data (i.e. factors of task allocation in GSD) is collected from published literature we will do a frequency analysis on it to know the frequently occurring critical factors. We will then validate our findings using online questionnaire given to industry experts. At the same time we aim to collect best practices to handle the identified challenges from industry with the help of our online questionnaire. This will help us in designing a comprehensive task allocation model for global software development. After that we implement task allocation model for real time project life cycle assessment (LCA) tool as well as one random data instance.

Our research methodology and approach can hence be summarized into the following phases:

Phase 1: Systematic Literature Review

In this phase, we aim to search and cover about 5 online research databases (i.e. IEEEXplore, ScienceDirect, ACM, John Wiley and Springer Link) for our SLR.

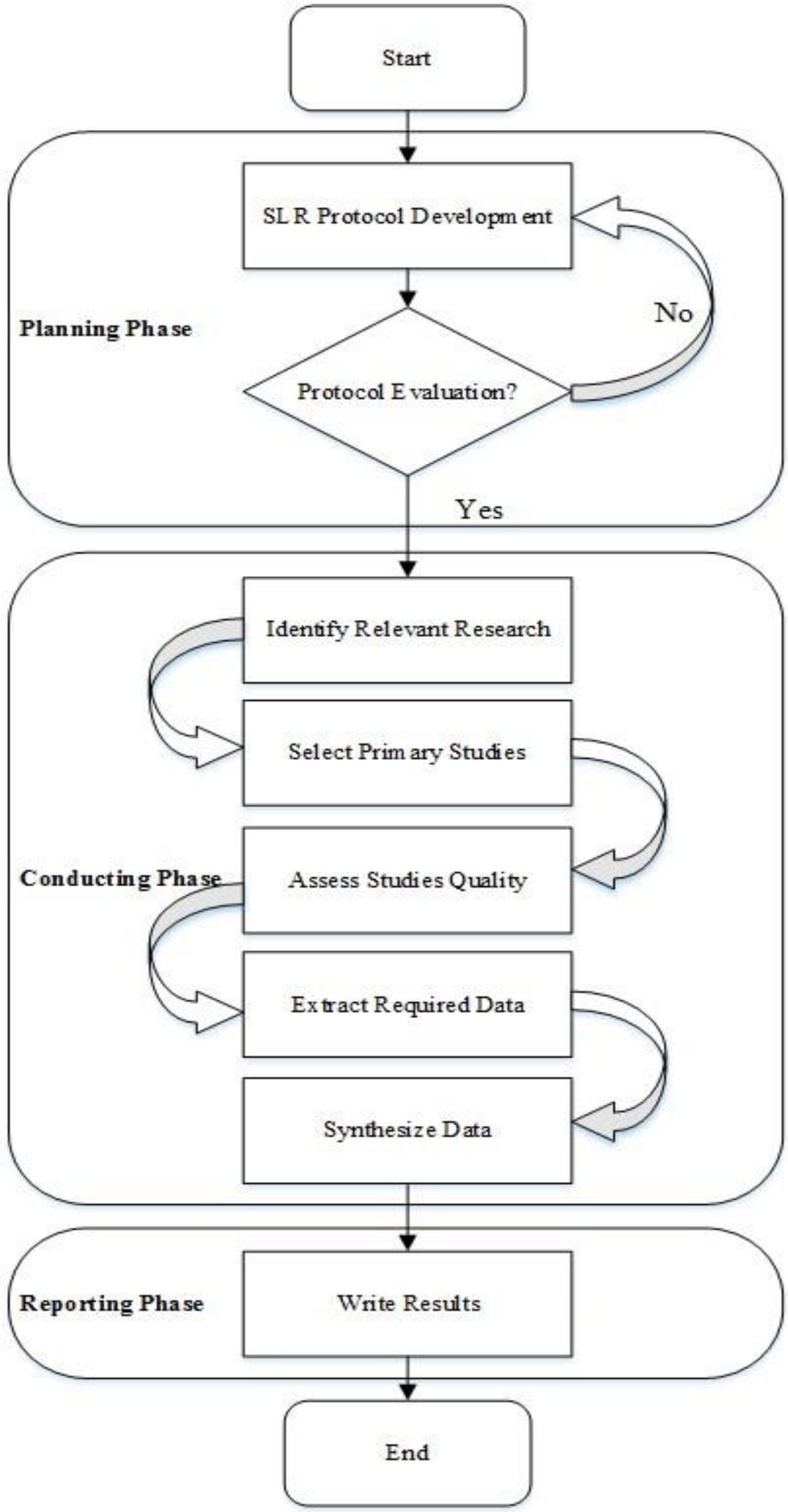

Figure 1.2 A Step by Step process of SLR

Phase 2: Empirical Study with Industry Experts

In this phase, we plan to validate our findings with the help of an online questionnaire given to industry experts.

Phase 3: Compare SLR findings with Industrial data

Compare SLR findings with industrial expert's data in order to find best match of industry with literature.

Phase 4: Implementation of Task Allocation Model

We implement task allocation model in GSD using genetic algorithm based on SLR and industrial expert's data.

Phase 5: Conclusions

The conclusion of the research is then presented.

1.5. Thesis Outline

The remaining sections of the thesis are organized as follows. Chapter 2 presents basic terminology and background information on GSD. We reviewed the related works in Chapter 3. Chapter 4 presents our two-phased research methodology. In Chapter 5, we present an in-depth analysis of our results. Chapter 6 covers the implementation of task allocation model using genetic algorithm based on empirical study results. Chapter 7 discuss the limitations of this work. Finally, Chapter 8 concludes the thesis and suggests some future work.

CHAPTER 2

BACKGROUND AND OVERVIEW

This chapter presents basic terminology and background information on GSD. Section 2.1 explains about GSD and various types of GSD.

2.1. What is Global Software Development?

Global software development, or software development outsourcing, is a recent software engineering paradigm which aims to develop quality software at low development cost[22]. Software development outsourcing is a relationship between client and vendor organizations in which a client contracts out all or part of its software development activities to one or more vendor, who provide agreed services in return for financial compensation [23]. Figure 2.1 depict the general overview of Global Software Development (GSD).

Figure 2.1 Global Software Development Overview

Different types of software outsourcing can be grouped into the follow two categories. Figure 2.2 presents the various types of outsourcing [24].

(i). Types of software outsourcing on the basis of geographic location:

On the basis of geographic distance between vendors and clients, outsourcing is categorized into three types: onshore outsourcing, near shore outsourcing and offshore outsourcing [25].

- **Onshore outsourcing**

Onshore outsourcing is also called domestic outsourcing, which consists of both domestic vendors and domestic clients [26]. This means that both (vendor and client) organizations are positioned in the same country.

- **Near shore outsourcing**

Near shore outsourcing or simply near shoring is defined as the transfer of software development work to a nearby foreign country to reap lower labor cost advantages [27]. The term Near shore was first introduced in a story about an entrepreneurial software development venture called PRT that was established in the Caribbean island of Barbados during the years 1995-1998[28]. During this period the word "near" referred to closeness to the United States from geographic point of view while "far" referred to the geographic distance of the client firms in the United States from the Indian vendors. An example of the Near shore outsourcing destination for the outsourcers in the United States is Canada [29].

- **Offshore outsourcing**

Offshore software development outsourcing refers to outsourcing in a geographically distant country. It is also referred to as far shore outsourcing in the published literature but the terms 'offshore outsourcing' or simply 'software outsourcing' have been used more frequently in the literature. The offshore activities have been going on from the past decade and are increasing quickly [30]. The major vendor countries for offshore outsourcing are China, Russia, Ireland and India whereas the client countries are the North America, Australia and Japan [31]. In providing offshore outsourcing services, India has a majority of the IT market share which is then followed by China [32].

(ii).Types of outsourcing on the basis of relationship

Oh and Gallivan [33] have categorized the offshore outsourcing relationships into four categories based on client and vendor contract. These are "Complex Relationships", "Co-Sourcing Relationships", "Multi-Vendors Relationships", and "Simple Dyadic Relationships".

- **Simple Dyadic Outsourcing Relationship**

Simple Dyadic is a one to one relationship between client and vendor. The client outsources its software to a single vendor who is responsible for the fulfillment of the job as per the clients' instructions [33].

In the case of a simple dyadic outsourcing relationship, when the relationship between vendor and client is on micro/personal level instead of a macro/organizational level, the outsourcing relationship is called Microsourcing [34].

Microsourcing is also termed 'personal work outsourcing', which is a type of outsourcing relationship on micro/individual level [34]. The situation occurs when an individual (client) outsources his/her own personal software development work to another person/programmer who provides services in return for financial compensation [34].

- **Multi-Vendors Outsourcing Relationship**

Multi-Vendors is one to many relationship between client and vendor. In this type of agreement/contract one client and many vendors are involved who consult each other to benefit from each other's' expertise and to settle the outsourcing task jointly [33].

- **Co-sourcing Relationship**

Co-sourcing is a many to one relationship between client and vendor (inverse of Multi-vendor Relationship). This type of relationship typically fulfill the common requirement of more than one clients. They therefore jointly outsource the software development project to a single vendor. This type of situation arises in organizations like hospitals etc. that need identical or similar software for some of their activities [33].

- **Complex Outsourcing Relationship**

This type of relationship comprises multiple clients and multiple vendors. This type of relationship typically fulfil the requirements of more than two client organizations when they need common software for their operations. [33].

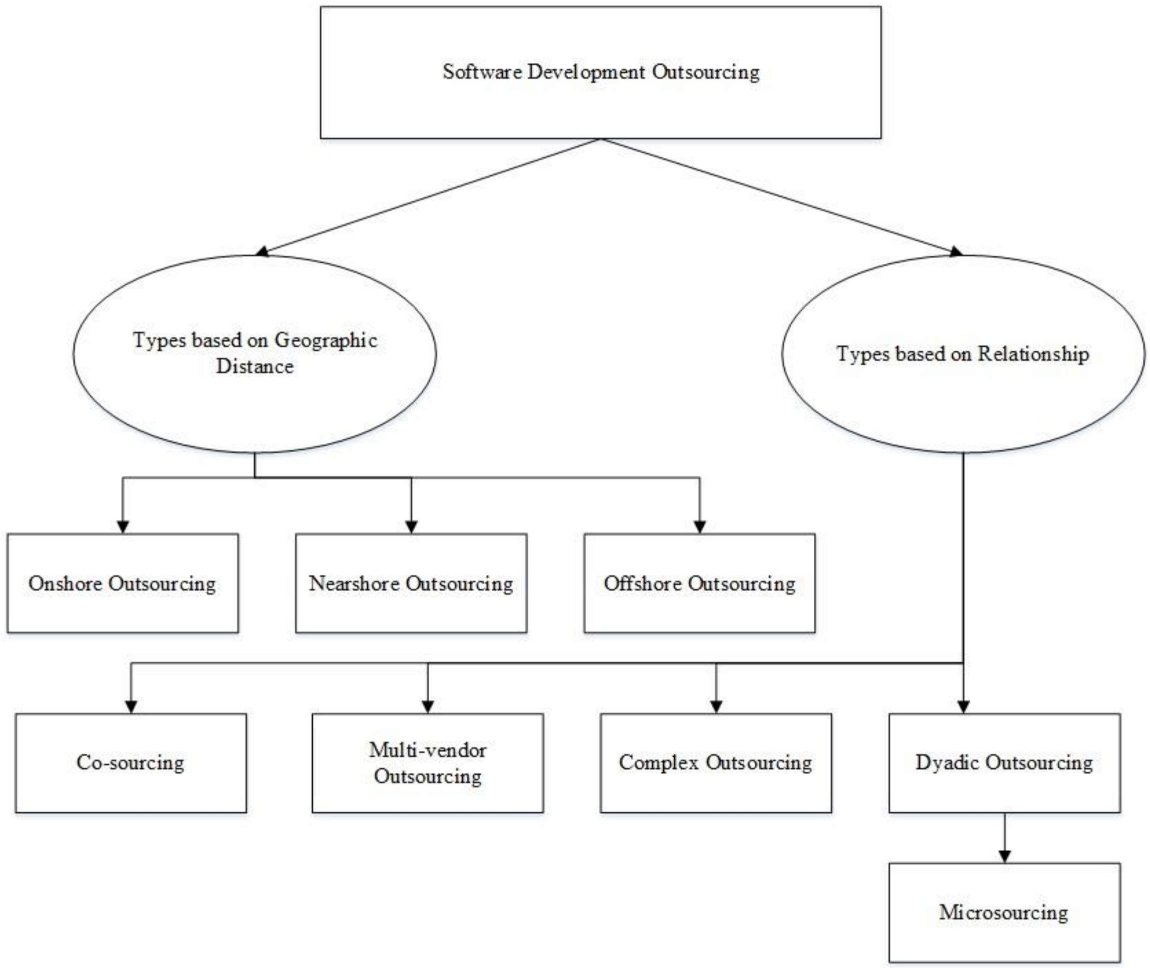

Figure 2.2 Types of Outsourcing

CHAPTER 3

LITERATURE REVIEW

In this section we present a brief review of the related and current literature with respect to motivation, risks and existing work on global software development. Motivations and risks of Global Software Development (GSD) are discussed in section 3.1. Section 3.2 gives an overview of the existing work carried out so far on GSD topic.

3.1. Motivation and Risks of GSD

There are several causes for initiating global software development project [27, 35]. Client organizations benefit from offshore outsourcing because typically cost 1/3rd less than onshore vendors and even less when compared with in-house operations in developing countries [36]. Amongst many other reasons for outsourcing, generally client organizations outsource their software development work to offshore locations to gain cost and quality advantages, access to leading-edge technology and the ability to focus on core competencies [27]. Moreover, offshore vendors improvise on their skills and quality of service with the increase in experience of offshore outsourcing projects [26]. Conversely quite apart from the outsourcing benefits there are many risks in an outsourcing process [9],[37], such as temporal incompatibility, cultural differences and hidden costs.

Reduction of cost is the major promoter for software outsourcing [11]. Other promoters for outsourcing comprises of access to cutting edge technology and to focus on core business model of the organization [38].

Offshore outsourcing is not a risk free activity as significant outsourcing failures have been reported [39]. Islam et al, [4]argue that lack of understanding between the client and vendor organization, ambiguous requirements and ineffective development processes may yield substantial risks. The results of a survey shows that eight out of every ten firms that have outsourced their software development project to an offshore vendor have faced major problems due to insufficient preparation and poor management by both the vendor organizations. King [40] reports that JP Morgan, a world renowned financial firm decided to go for in house software development which lead to non-renewal of USD 5 billion $ contract with IBM.

3.2. Existing Work on Task Allocation in GSD

This section presents a review of the key studies conducted on the topic task allocation in GSD environment. The objective is to summarize and discuss the results of each study in order to better recognize the context of the problem.

Hanssen et al. [41] presented a systematic literature review study with a focus on the application of agile methodologies in GSD. They analyzed 12 SLRs and they reported that agile practices is critical factor for the successes of distributed projects. Richardson et al. [42] carried out study based on three industrial case studies and three literature review to identify factors which are important to global software engineering.

They also used to the literature and empirical data to identify threats to software projects due to lack of implementation of global software engineering team processes.

Lately, Marques et al. [43] presented a tertiary study to categorize systematic reviews conducted in distributed software development context. The tertiary study identified fourteen SLRs addressing different aspect of distributed software development projects. Seven SLRs studied different aspects of managing distributed development. Furthermore, four SLRs addressed topics of engineering process and three SLRs are related to requirements and design issues of GSD projects. However, Marques et al. [43] concluded that topics covered by existing SLRs in GSD are limited and amount of empirical studies is relatively small.

Kroll et al. [44] presented Systematic literature review (SLR) to address the challenges and best practices for follow the sun development process. Follow the sun is subclass of Global software development in which software development life cycle is spanned over 24 hour in order to decrease the total development time [45]. The study explore challenges like communication (Language difference, communication difficulties etc.), coordination (time zone difference, number of sites, geographical differences etc.) and culture (cultural differences, different technical backgrounds). This study also discuss Best practices as findings of SLR like time window, agile methods, time zone management, pair programming and proper task distribution. However, Kroll et al. [44] discuss task allocation as best practice but not how we can do this best practice. He also concluded that there is still a gap of GSD techniques for follow the sun development.

Carmel et al. [45] presented mapping study to explore the challenges and best practices for project management in global software development paradigm. This study results explore different challenges like time zone difference, cultural difference, communication problem, task allocation across development sites. As Carmel et al. [45] present task allocation as a challenge for project management and project can be delayed or over cost. However, Carmel et al. [45] does not discuss how we can address this challenge, thus it needs more investigation.

Mohapatara et al. [46]presented systematic literature review (SLR) to address the task allocation challenges in global software development process. This study discusses the issue of task allocation with respect to roles and responsibilities. Role is particular designation of someone in the organization and his behavior becomes the responsibilities. [46] find the factors that can effect task allocation in global software development like (time zone difference, cost, communication & coordination, task dependency etc.) and then develop a model that relate the roles and responsibilities to address these factors that can affect. This study build model to distribute roles and responsibilities to address the task allocation issues and can be viewed as human resource distribution to tasks. However, [46] miss the site structure that is another perspective of task allocation and initiate challenges in task allocation in GSD.

Lamersdorf et al. [47] presented an interview based qualitative study aimed at identifying and understanding different criteria used in practice. The study shows that the sourcing strategy and the type of software to be developed have a significant effect on the applied criteria. The main task allocation criteria are labor costs, proximity to market, turnover rate and strategic planning.

Lamersdorf et al. [48] presented a risk driven customizable model to suggest a set of task allocation alternatives based on project and specific characteristics and analyzing it with respect to potential project risks related to work distribution. Furthermore, they evaluated the proposed model by a series of semi-structured interviews in a multination IT organization.

Narendra et al. [49] presented an integrated formal technique to develop task allocation model for GSD projects. The proposed approach generates effort estimation for the new task allocation based of factors such as expected general percentage allocation of overall effort estimates to each task and effort estimates for executing a task at a particular site.

Wickramaarachchi and Lai [50] proposed a method for work distribution to different locations with an aim to minimize overhead costs. The method categorizes the offshore tasks based on software process model. It also proposes a method to distribute work to suitable tasks using work specific matrix, work dependency Matrix and site dependency matrix.

In summary, GSD continues to experience substantial growth and is different to traditional single site based development[42]. GSD project managers need support in successfully managing task allocation, an important component of the project management body of knowledge, in a global environment. However, traditional task allocation techniques do not explicitly cater for the complex needs of GSD projects. This draws our attention to a number of potentially interesting results which we are going to explore in this research work. It is evident that the study of task allocation and work distribution in GSD is emerging research trend both in academia and industry, however, it underlines the need for further research.

CHAPTER 4

RESEARCH METHODOLOGY

4.1. Introduction

We followed a two-phase approach in making our research a comprehensive study. In order to address our research questions, we applied the Systematic Literature Review (SLR) and empirical survey approaches. In the first phase we determined the challenges via a Systematic Literature Review (SLR). We then complemented the findings with a Questionnaire survey. Figure 4.1 depict the complete research methodology.

We discuss the each of the research methodologies in detail in the following sections. Section 4.2 explains the whole SLR process which includes developing an SLR protocol, clean and processes the findings via initial and final study selection, validation and filtration using quality assessment techniques, data synthesis and proof reading. Section 4.3 explains the Questionnaire Survey in detail which was answered by 41 experts belonging to Fortune 500 companies and various geographical locations across the globe. The participants were asked to rank each challenge on a five-point scale to determine the perceived importance of each challenge. The survey included challenges identified in during our systematic literature review.

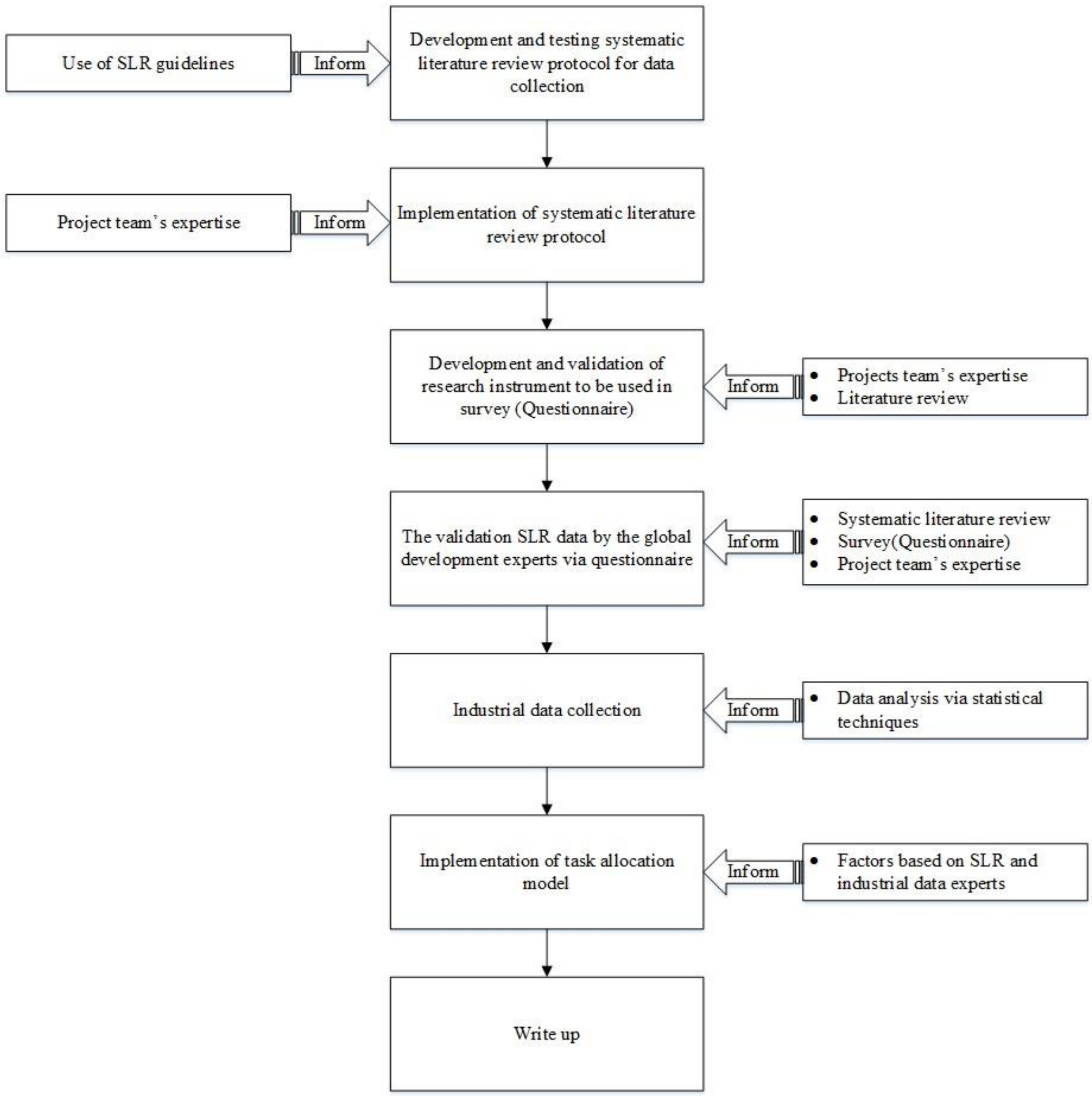

Figure 4.1 Research Methodology

4.2. Data Collection and Analysis via SLR

"A SLR is a defined and methodical way of identifying, assessing, and analyzing published primary studies in order to investigate a specific research question"[21, 51]. SLRs is a formally planned activity to address literature that make it different from ordinary literature surveys. In finding, evaluating, and summarizing all available evidence in the form of studies on global software development, a SLR may provide a greater level of validity in its findings than might be possible in any one of the individual studies surveyed in the SLR.

Keele [21] explains SLR with the help of three main phases (1) “planning the review”, “conducting the review”, and “reporting the review”. Each of these phases contains a sequence of stages, but the execution of the overall process involves iteration, feedback, and refinement of the defined process.

The output from the planning phase is a systematic review protocol that defines the purpose and procedures for the review. A systematic review protocol is a formal and rather concrete plan for the execution of the systematic review. Kitchenham notes that a pre-defined protocol is necessary to reduce the possibility of researcher bias. The contents of a systematic review protocol in many ways foreshadow the structure of the final report it describes the background context for the research, the specific research questions, the planned search strategy, criteria for publication selection, the treatment of publication quality assessment, the data extraction plan, the data synthesis plan, and a project plan. This conducting phase ultimately generates final results, but also generates the following intermediate artifacts: the initial search record and archive, the list of selected publications, records of quality assessments, and extracted data for each of the selected publications. In this phase the selection process is intended to identify the found primary studies that provide direct evidence about the research questions. Again, the selection process should follow the plan described in the protocol. Kitchenham describes selection as a multistage process: first researchers only exclude clearly irrelevant publications; and then from the resulting short list researchers only include publications that contain extractable data addressing the research questions. After selecting the primary studies, a more detailed quality assessment is needed to allow researchers to assess differences in the implementation of studies. For detailed quality assessment, checklists can be designed using factors that could bias study results.

In order to conduct the SLR process five digital libraries are used as shown in Figure 4.2 (Based on the available access): Since these libraries differ in their search mechanism and capability, we tailored our search strings accordingly. Complete SLR protocol is attached in appendix A.A.

Inclusion and exclusion for primary studies are shown in Figure 4.3.

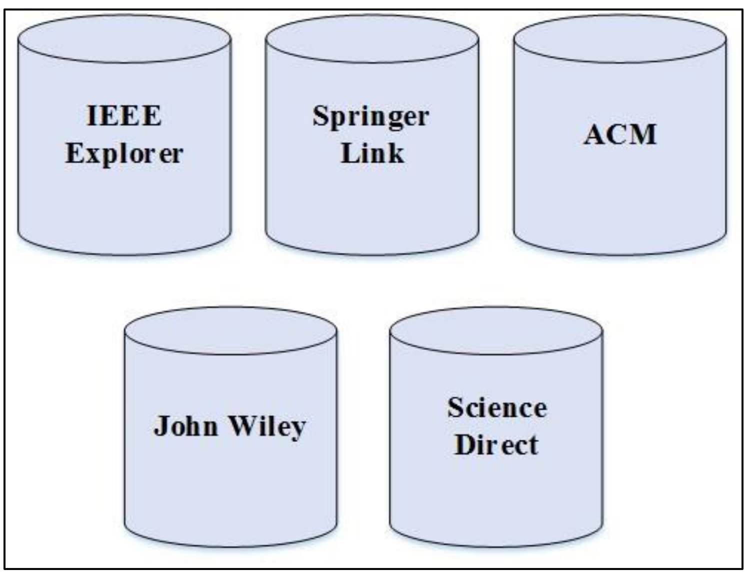

Figure 4.2 Digital Libraries used in the Work

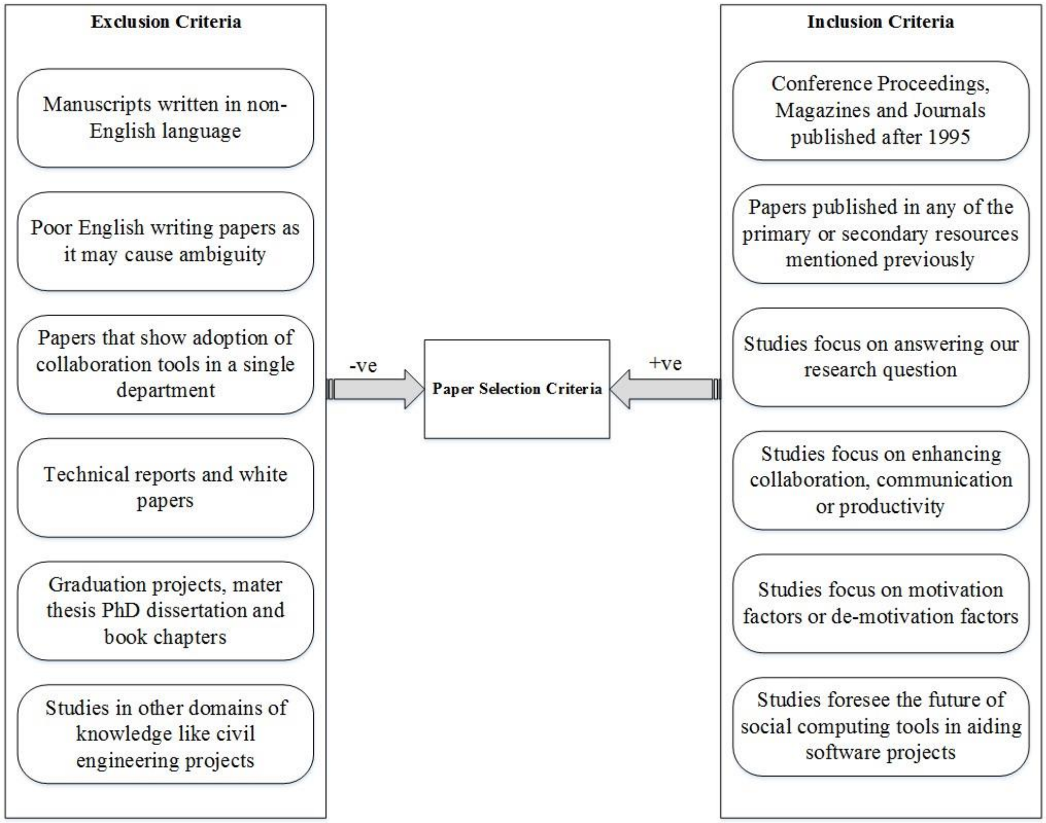

Figure 4.3 Inclusion and Exclusion Criteria for Primary Studies

At the end of SLR process 38 studies has been selected for data extraction. Figure 4.4 shows the complete SLR process results.

4.2.1. Quality Assessment

For any paper to pass the initial phase, a quality assessment was done. We have to assess the quality of the literature selected after final selection for its quality. The quality assessment activity for the relevant literature will be carried out at the same time during the extraction of relevant data so as to ensure that a valuable contribution is made to the SLR. We will detail a quality assessment checklist that will provide means to quantitatively assess the quality of the evidence presented by these studies. However, these checklists are not meant to be a form of criticism of any researchers' work and any changes to the quality assessment criteria as such will be documented. Nine quality criteria were prepared as shown in Table 4.1.

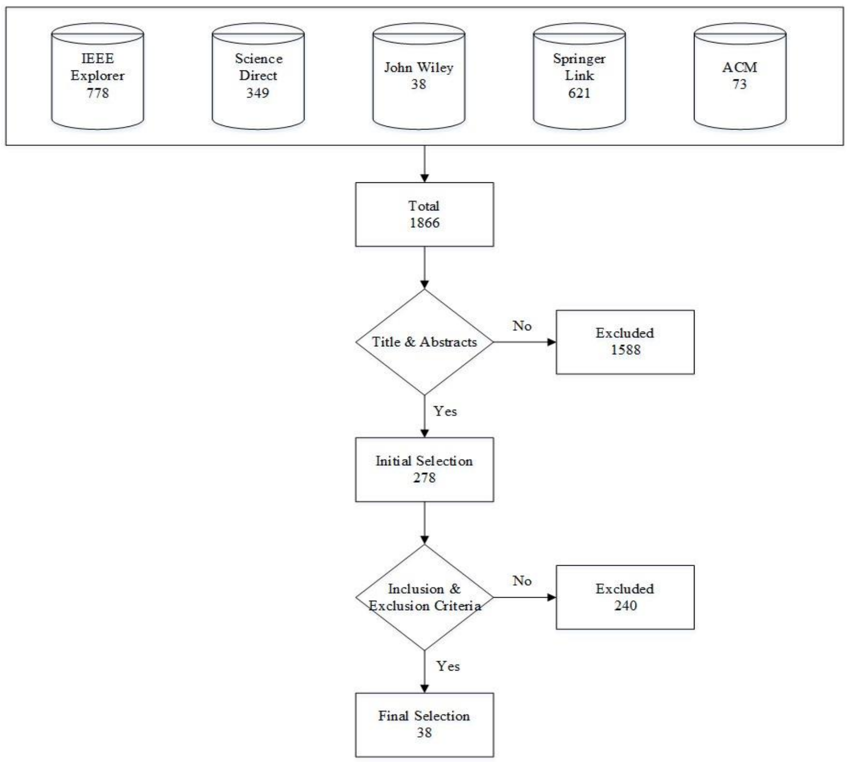

Figure 4.4 SLR Process Implementation

Table 4.1 Quality Assessment Criteria

#	Questions	Possible Answers		
1	Is there a rationale for why the study was undertaken?[52]	Y=1	N=0	P=0.5
2	Is the paper based on research or report based on expert opinion?[53]	Y=1	N=0	
3	Is the research goals are clearly reported?[53]	Y=1	N=0	P=0.5
4	Is the proposed technique clearly described?[54]	Y=1	N=0	P=0.5
5	Is the research supported by a tool?[55]	Y=1	N=0	
6	Was the research empirically validated?[55]	Y=1	N=0	
7	Is the research results clearly described?[55]	Y=1	N=0	P=0.5
8	Is there is explicit discussion about the limitations of this research?[56]	Y=1	N=0	P=0.5
9	Does study suggest solutions for task allocation in GSD?	Y=1	N=0	P=0.5

Each paper evaluated against this quality criteria. Eight out of nine quality criteria questions were adopted from existing literature and one is applied based on this study.Q1,Q2,Q3,Q4,Q5,Q6,Q7,Q8 were adopted from literature and Q9 proposed based on the scope of this study. In order to define quality criteria we relied on SLR's published in reputable journals in the area of empirical software engineering. Q1 was adopted from [52], Q2 and Q3 were adopted from [53], Q4 was adopted from [54], Q8 from [56] and Q5, Q6 and Q7 from [55].

Q2, Q5 and Q6 scores were determined using the two-grade scale score (yes/No). If any study fulfill required criteria, the study received 1 point in this question, otherwise it received 0 point in this question. Q1, Q3, Q4, Q5, Q7, Q8 and Q9 scores were determined based on three grade scale score. If the study full fill the required criteria then study will receive 1 point but on the other hand study full fill the required criteria but not too much clear or not strong enough the study will receive 0.5 otherwise study will receive 0 point in this question.

After the final selection of primary studies depending upon the quality assessment criteria we have to start with the data extraction phase of the systematic literature review process. We will use the data extraction form to extract the data. The data will be extracted by a single reviewer and will be assessed by a PhD supervisor in a random manner. Table 4.2 represents the data extraction form which will be used for the purpose of extracting relevant data from primary studies:

Table 4.2 Data Extraction Form

Data Item	Value	Supplementary Notes
Study Information Data		
Paper ID		
Date of Review		
Title		
Author(s)		
Year of publication		
Geographical Location		
Reference type	Journal/Conference/Thesis/Unpublished	
Type of Study	SLR/Interview/Case Study/Report/Survey	
Publisher		
Perspective	Client/Vendor	
Techniques/Algorithm		
Tool Support	Yes/No	
Tool	Developed/Prototype	
Data Relevant to Answering Research Questions		
Factors that influence Task allocation in GSD		
Solutions to the identified factors in GSD		

The data extracted from the primary studies will be saved as a Microsoft Excel document in < paper id >_<author name>_<year of publication> format. After the extraction of data we will use the data synthesis form as shown in

Table 4.3, to summarize and compile the extracted data from the primary studies so as to answer each of the research questions. This form helps to carry out various types of statistical analysis so as to draw conclusions. . As with other stages, the procedures to be followed should be defined in the protocol.

Table 4.3 Data Synthesis Form

RQ1: What are the factors that influence task allocation in GSD?							
Paper ID	Quality (score)	Population (e.g. task allocation)	Geographical Location	Year of study	Perspective	Type of Study	Factors that influence task allocation

Reporting the review is a single stage phase. Usually, systematic reviews are reported using two formats: in a technical report and in a journal or conference papers. The structure and contents of reports is presented in the guidelines[21].

4.3. Statistical Analysis

In order to analyze the SLR data we will use frequency analysis, because frequency analyze are helpful for treatment of descriptive information. Number of occurrences and percentages of each data variable can then be reported using these frequency tables. Frequencies are helpful for comparing and contrasting within groups of variables or across groups of variables and can be used for both nominal/ordinal as well as numeric data.

Coding in empirical research is one method of extracting quantitative data from qualitative data in order to perform some statistical analysis[57]. In this research data from the literature will be categorized and coded in order to perform frequency analysis. We will measure the occurrence of key items in a survey of the literature. By comparing the occurrences of key items in a number of articles against occurrences of other key items in the same articles, we will calculate the relative importance of each item. For example, a percentage of x for item y means that item y is mentioned in x% of the literature, i.e. if an item is mentioned in 10 out of 20 articles, it has an importance of 50% for comparative purposes. In addition to frequency analysis we will also use rigorous statistical techniques (Chi-Square Test, Spearman's Rank order Correlation etc.) in order to quantify the collected data. In this way we will compare and rank different items.

4.4. Data Collection and Analyzes via Survey (Questionnaire)

The data identified through the SLR will be validated by the global software development experts via questionnaire. Industrial surveys will be conducted in this questionnaire in order to seek experiences of global software development experts about the findings identified through SLR process. The GSD experts will be selected on the basis of their experience in the field of global project management. It is also important to acknowledge that the project will solicit information regarding the experiences of those experts who are tackling real GSD issues on a daily basis.

Content analysis is a research method that uses a set of procedures to make valid inferences from text[58]. Various authors provided definitions of content analysis. For example, [59] state, "content analysis is any research technique for making inferences by systematically and objectively identifying specified characteristics within text".

Once data transcripts have been collected, the next task is to develop categories to be used in comparing and contrasting results. Identification of categories is one of the important stages in analyzing qualitative data[60]. Different researchers describe different approaches for establishing categories for qualitative data[61, 62].

This research seeks to identify key criteria and experiences of practitioners about task allocation in GSD projects. The following process will be used to analyze the data:

- Data collection from questionnaire: All the questionnaire transcripts will be read to identify the major task allocation techniques and best industrial practices.
- Generate categories: All the questionnaire transcripts will be read again to generate categories for responses.
- We will identify relation within and between data categories.
- We will interpret the categorized data and their relationships in order to identify key points and lessons learnt.

4.5. Statistical Analysis to Compare SLR and Questionnaire Data

This section presents a comparative analysis of the factors identified by the SLR and the questionnaire survey. This will help in understanding the similarities and differences among the outcomes of the two data sets (Literature findings and industrial data).

In order to quantify the significance of the similarity in the factors identified using the SLR and the questionnaire-based empirical study, we will use frequency analysis and other statistical techniques[57, 63](Chi-Square Test, Spearman's Rank order Correlation etc.) to quantify the collected data.

CHAPTER 5

EMPIRICAL STUDY RESULTS AND DISCUSSION

5.1. Introduction

In this chapter we present the results and analysis of our study. Section 5.2 explains the quality assessment of primary studies. Section 5.3 explains the research model which include findings from the SLR, findings from Questionnaire, client vendor analysis of SLR findings as well as industrial expert's data and also compare the both results using various statistical techniques like Chi-Square Test, Fisher Exact test and T- Test.

5.2. Quality Assessment Results

The accuracy of reliability of data extraction results can be increased by applying proper quality assessment criteria. Primary studies of this research are evaluated by using the quality criteria explained in Section 3.5. This evaluation helped to determine the validity of the implications and expressive synthesis of results.

The quality assessment results for each study are shown in Table 5.1. The acceptable threshold for each study is set 50% with the aim to improve the quality of our results. Regarding the implications of quality criteria, all studies are more than 50% and average score is 6.4. Q1, Q2, Q3 and Q4 received average score more than 90% which shows that each study fulfil the basic criteria of research goals. Q6, Q7 and Q8 received average score more than 50%, however Q5 and Q9 got less than 50% which implies that these two most of the literature discuss the problem but did not provide the solutions for these problems. One other point that needs consideration in this area is the lack of tool usage, and this is directly related to the first point proposed solutions. However overall quality of all primary are acceptable.

Table 5.1 Quality Assessment Results

ID	Author	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Total Score	Qual. (%)
1	Lamersdorf et al. [64]	1.0	1.0	1.0	1.0	0.0	1.0	0.5	0.5	0.0	6.0	66.7
2	Lamersdorf et al. [65]	1.0	1.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	7.0	77.8
3	Lamersdorf et al. [66]	0.5	1.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	6.5	72.2
4	Munch et al. [67]	1.0	1.0	1.0	1.0	0.0	1.0	0.5	0.5	0.0	6.0	66.7
5	Pereira et al. [68]	1.0	1.0	1.0	0.5	0.0	0.0	1.0	0.5	0.0	5.0	55.5
6	Nakakoji et al. [69]	1.0	1.0	1.0	1.0	0.0	1.0	1.0	0.5	0.0	6.5	72.2
7	Hu et al. [70]	1.0	1.0	1.0	1.0	0.0	1.0	0.5	0.5	1.0	7.0	77.8
8	Gupta et al. [71]	1.0	1.0	1.0	1.0	1.0	0.0	1.0	0.0	0.0	6.0	66.7
9	Vathsavayi et al. [72]	1.0	1.0	1.0	1.0	0.0	1.0	0.5	0.0	0.0	5.5	61.1
10	Lamersdorf et al. [73]	1.0	1.0	1.0	1.0	1.0	1.0	0.5	0.5	0.0	7.0	77.8
11	Helming et al. [74]	1.0	1.0	1.0	1.0	0.0	1.0	0.5	0.5	0.0	6.0	66.7
12	Yilmaz et al. [75]	1.0	1.0	1.0	1.0	0.0	1.0	1.0	0.5	0.0	6.5	72.2
13	Doma et al. [76]	1.0	1.0	1.0	1.0	1.0	1.0	0.5	0.5	0.0	7.0	77.8
14	Celik et al.[77]	1.0	1.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	7.0	77.8
15	Jalote et al. [78]	1.0	1.0	1.0	1.0	0.0	1.0	1.0	0.5	0.0	6.5	72.2
16	Barcus et al.[79]	1.0	1.0	1.0	1.0`	0.0	0.0	1.0	0.5	1.0	6.5	72.2
17	Abufardeh et al. [80]	1.0	1.0	1.0	0.5	0.0	0.0	0.5	0.5	0.0	4.5	50
18	Lamersdorf et al. [81]	1.0	1.0	1.0	1.0	0.0	1.0	1.0	0.5	0.0	6.5	72.2
19	Setamanit et al. [82]	1.0	1.0	1.0	1.0	1.0	1.0	0.5	0.5	1.0	8.0	88.9
20	R. Mayoral et al. [83]	1.0	1.0	1.0	1.0	0.0	1.0	1.0	0.5	0.0	6.5	72.2
21	Wickrammaarachchi et al. [50]	1.0	1.0	1.0	1.0	0.0	1.0	0.5	0.5	0.0	6.0	66.7
22	Lamersdorf et al. [48]	1.0	1.0	1.0	0.5	0.0	1.0	1.0	0.5	0.0	6.0	66.7
23	Mak et al. [84]	1.0	1.0	1.0	0.5	1.0	1.0	0.5	0.5	1.0	7.5	83.3
24	Lamersdorf et al. [85]	1.0	1.0	1.0	1.0	0.0	1.0	0.5	0.5	0.0	6.0	66.7
25	Narendra et al. [49]	1.0	1.0	1.0	0.5	0.0	1.0	1.0	0.5	0.0	6.0	66.7
26	Marques et al. [86]	1.0	1.0	1.0	1.0	0.0	1.0	1.0	0.5	0.0	6.5	72.2
27	Lamersdorf et al. [47]	1.0	1.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	7.0	77.8
28	Cataldo et al. [87]	1.0	1.0	1.0	1.0	0.0	1.0	0.5	1.0	0.0	6.5	72.2
29	Fernandez et al. [88]	1.0	1.0	0.5	0.5	0.0	0.0	1.0	0.5	0.0	4.5	50
30	Mak et al. [89]	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.5	0.0	7.5	83.3
31	R. Mayoral et al. [90]	1.0	1.0	1.0	1.0	0.0	1.0	0.5	0.5	0.0	6.0	66.7
32	Setamanit et al. [91]	1.0	1.0	1.0	1.0	1.0	0.0	1.0	0.0	0.0	6.0	66.7
33	Richardson et al. [92]	1.0	1.0	1.0	1.0	0.0	1.0	0.5	1.0	0.0	6.5	72.2

34	Imtiaz et al. [93]	1.0	1.0	1.0	1.0	0.0	1.0	1.0	0.5	0.0	6.5	72.2
35	Deshpande et al. [94]	1.0	1.0	1.0	1.0	0.0	1.0	1.0	0.5	0.0	6.5	72.2
36	Lamersdorf et al. [95]	1.0	1.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	7.0	77.8
37	Battin et al. [96]	1.0	1.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	7.0	77.8
38	Goldman et al. [97]	1.0	1.0	1.0	1.0	0.0	0.0	1.0	0.5	1.0	6.5	72.2
Average		0.98	1.0	0.9	0.92	0.18	0.8	0.76	0.57	0.1	6.40	

5.2.1. Overview of Studies

This section explain the overview of studies from temporal point of view. Papers published between 1995 and 2014 are reviewed in this study. Figure 5.1 Temporal View of the primary studiesFigure 5.1 reveals that most of the papers during 2009 to 2012 (63%). In 2009 (18%), 2010, 2011 (16%) and 2012(13%) but after that there is drop in this area and also there is no trend in this topic. It is also worth noting that in 2013 there are only 3 studies and in 2014 only 1 studies, which shows researcher loose interest in this area but there are lot of problem still unsolved.

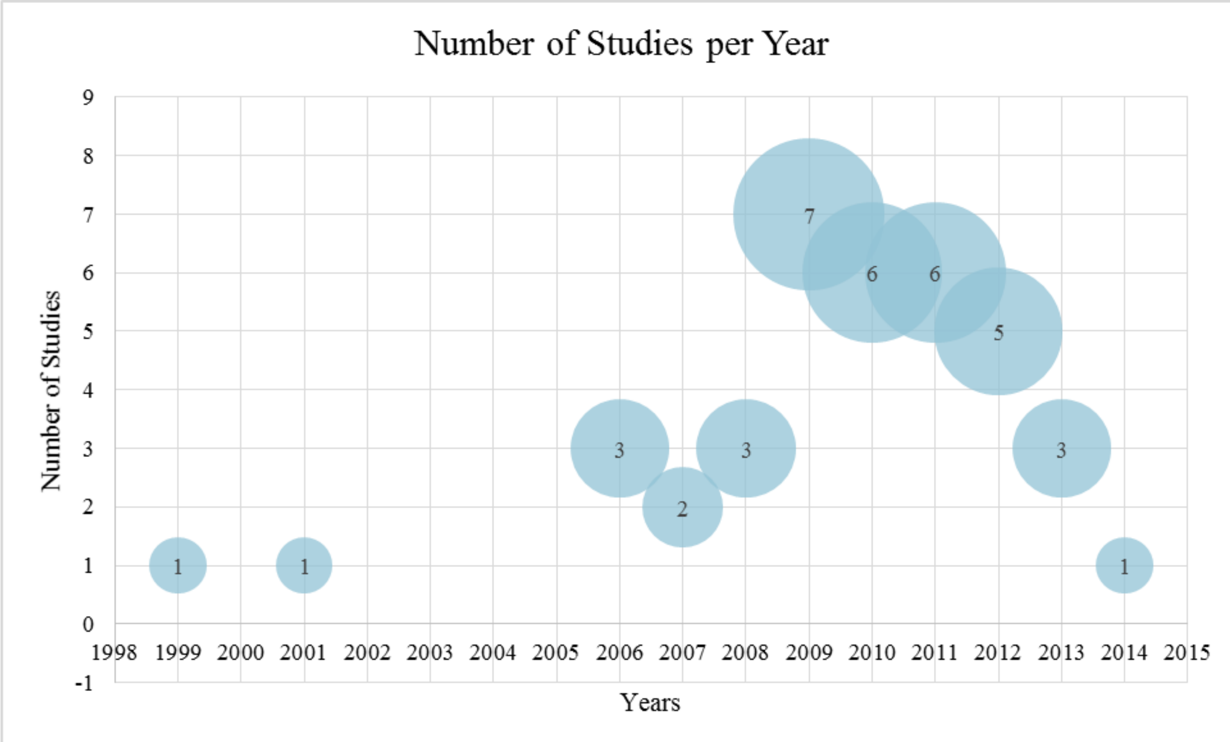

Figure 5.1 Temporal View of the primary studies

5.3. Empirical Study Research Model

This section present the literatures findings and industrial analysis based on client vendor by using statistical technique (Chi-Square test, Fisher Exact test) and also compare the results of literature with industrial experts' opinion using statistical technique.

In order to analyze data based on client vendor perspective two hypothesis are designed. One is null and other is alternative hypothesis. All factors are analyzed against these two hypothesis individually. Hypothesis acceptance OR rejection is based on P-Value and confidence level is set at 95%.

H1 (Null Hypothesis): There is an association between client and vendor perspective about particular task allocation factor OR they are equal.

H2 (Alternative Hypothesis): There is no association between client and vendor perspective about particular task allocation factor OR they are not equal.

Primary studies are categories into two perspectives client and vendor. The distinction is based on either the explicit or based on the region where the study is undertaken. After carefully reading each selected paper, a client-vendor categorization was made for the identified factors. Table 4 shows the client and vendor countries where research was conducted for the papers included in our SLR study. A total of 23 studies were carried out in client countries while 15 studies were conducted in vendor countries.

Table 5.2 Studies Count based on Client Vendor Countries

Client Countries	Count	Vendor Countries	Count
Australia	1	Brazil	2
USA	6	China	2
Ireland	2	India	4
Finland	1	Turkey	1
Germany	9	Spain	5
Japan	1	Pakistan	1
UK	3		
Total	23	Total	15

5.3.1. Findings from SLR

This section presents the initial SLR-based literature survey results. The total number of results retrieved after inputting the search terms in the electronic databases are shown in Figure 4.4. After the initial round of screening by reading the title and abstract, about 278 studies belonging to five different electronic research databases were selected. After full text readings in the second screening and application of inclusion and exclusion criteria, about 38 primary studies were finally selected. To answer the research question, the data was carefully extracted and synthesized from the thirty eight finally selected studies. We have identified a list of eleven criteria that influence task allocation process in global software development, is shown in Table 5.3.

Table 5.3 List of Task Allocation Factors

Factors	Frequency (No. of papers (n)=38)	Percentage
Site technical expertise	26	68
Time zone difference	24	63
Resource cost	18	47
Task dependency	17	44
Vendor reliability	14	36
Task size	11	29
Vendor maturity level	8	21
Local government regulations	5	13
Requirements Stability	3	7
Product architecture	3	7
Intellectual property ownership	1	2

In our study, the most highly cited criterion for task allocation in GSD projects is ‘site technical expertise’ (68%). The development sites are spread across geographical boundaries and each site has particular expertise i.e. Programming skills, tool usage skills that influence the product quality as well as other factor that impact on project; hence, selecting sites with appropriate domain expertise and knowledge is crucial to the success of a GSD project. This factor mainly insure the product quality along with budget and time requirement. For example, the interview of GSD project managers [47] revealed that matching specific technical skill sets available at a vendor site plays is one of the most important criteria for task allocation.

The second highest frequently mentioned criterion is ‘time zone difference’ (63%). Lamersdorf et al. [47] argues that time zone difference have positive as well as negative impact on overall effort. GSD project managers typically use ‘time zone difference’ to their advantage and decrease the overall delay by allowing 24 hour development “follow the sun”[64], “round clock development” [73] under certain conditions like mature process and ultimately decrease overall effort. On the other hand time shift between sites increase the communication and coordination problems that leads towards increase delays and overall effort[82] and also time zone difference may leads toward night shift that decrease employee motivation and ultimately decrease productivity [48].

‘Resource cost’ is another key criterion (reported by 47% of the articles selected from the SLR) for work distribution in a GSD project. In general, researchers and practitioners report that resource cost consideration is an important factor during development of globally distributed projects. Typically, project managers aim to assign work units to low labor cost sites.

On the other hand, GSD practitioners have also highlighted that cost alone should not be used as a sole criterion for task allocation because highly coupled tasks assigned to different sites potentially contribute to increase in communication and project execution costs [75, 86]. Another factor that needs consideration for choosing low cost is related to required technical expertise on that site that directly impact on software quality [85]. For this factor there is tradeoff between cost and product quality and project manager needs to consider the interest factor (Quality OR low cost).

Task Dependency is another key criterion and more than 44 % of the articles mentioned ‘task dependency’ as an important criterion for work distribution decisions in globally distributed projects. Jalote et al. [78] argues that it increase the overall development time and also limit the benefits of having multiple sites/resources (as we cannot start next task until previous cannot be finished, So most resources becomes waste). This factor has positive as well as negative influence on many other factors like time zone [64, 73] and resource cost [75, 86] that needs consideration collectively in task allocation process.

Two other factors are related to vendor organization are ‘vendor reliability’ and ‘vendor maturity level’ with 36% and 21% respectively. The researchers have discuss both factors in terms of respective past experience provides an important insight during the task allocation process of a GSD project. [82, 91] explains them in term of member familiarity that impact team performance, more familiarity between team members better the performance of team. They also argue that distance between teams can negatively impact the organization performance as physical distance cause communication and coordination’s problems. However, this affect can be mitigated by number of factors i.e. [48] describe that mature process can overcome communication problems that ultimately effects productivity and [81] discuss that better CMMI level (over all site process) of site can overcome this problem and that leads toward better productivity. In summary researchers have cited that the perceived reliability of a particular vendor helps clients to better manage task allocation risks in global teams.

'Local government regulations' is another factors and 13% research cited this as a factor for task allocation in GSD projects. [50] describe political relation between countries as the measure of compatibility between GSD sites and [47] describes political reasons as the terms and conditions of country for labor force i.e. work can be assigned with in the country, working hour regulations, salaries rules etc.

Other key factors in work distribution through task allocation in GSD teams are 'requirements stability', 'product architecture' and 'intellectual property ownership'. They have been depicted in our study where 7%, 7% and 2 % of the articles have mentioned them as task allocation criteria in GSD projects, respectively. [81] discuss it as "Degree of change in the requirements during the project" and ultimately impact the overall effort overhead.Furthermore, less frequently mentioned factors are 'product architecture' and 'intellectual property ownership'.

5.3.2. Findings from Questionnaire survey

In the second step of our research, we developed an empirical study survey questionnaire based on the factors identified in the SLR-based literature study. Industry experts gave their opinion to answer our research questions. Table 5.4 shows the rankings of the factors identified from our empirical study. This explains the view of the industry practitioners to assess a particular factors for task allocation in GSD. The table has been divided into 2 columns, i.e. 'Positive', and 'Negative'. The values present in the 'Positive' column shows the percentage of respondents who agrees with the identified factors of task allocation in GSD. Whereas, the values present in the 'Negative' column shows the percentage of respondents who feel the factors might not be present during task allocation in GSD. In order to measure the relative importance between positive respond factors we also ask from practitioners to rank them on five level scale 'very important', 'important', 'moderately important', 'somewhat important', '' and 'not important'. At the end of the survey we asked the practitioners to provide best practices as per their vast experience to handle those factors successfully. These best practices will help us in developing task allocation model. The survey sample is attached in appendix A.B.

Table 5.4 Industrial Expert's Data

Factors	Organizations' Observation (No. of papers(n)=62)					
	Positive			Negative		
	Strongly Agree	Agree	% age	Disagree	Strongly Disagree	% age
Site Technical Expertise	26	35	98	1	0	2
Requirements Stability	35	25	97	1	1	3
Product Architecture	28	31	95	3	0	5
Task Dependency	36	23	95	3	0	5
Resource Cost	27	31	94	4	0	6
Intellectual Property Ownership	16	38	87	7	1	13
Task Size	22	32	87	8	0	13
Vendor Maturity Level	24	29	85	9	0	15
Vender Reliability	27	21	77	13	1	23
Time Zone	16	29	73	13	4	27
Local Government Regulations	24	18	68	18	2	32

Site technical expertise in the most positive respond factor from industrial experts received 98% positive response and only 1 respondent considered it as irrelevant. This factor is very important for task allocation for GSD projects and has direct impact on product quality as one of the respondent explains this as follows;

> *"We allocate tasks keeping in mind the expertise, amount of work load and ease of our employer".* Senior Software Engineer

The high priority of this factor for task allocation in GSD projects are also supported by the literature (22 studies out of 38) shown in Table 5.3.

The next ranked factor for task allocation from industrial experts is 'requirements stability' and received 97% positive response. Researcher discuss this as degree of change in requirements during the project life span. In this way this factor has direct impact on product quality as well as product stability. Project budget can be increased if there is lot of changes in requirements during the project development and other later phases. One of the responded says that;

> "*We use proper requirements modeling techniques, requirements prioritization for GSD projects that decrease the overall effort".* Team Lead

However literature does not support the industrial expert's argument and only 7% studies list it as task allocation factor for GSD projects.

'Product architecture' is next ranked factor of task allocation in GSD projects received 95% positive response. This factor can be considered with requirements stability collectively as stable requirements result is stable architecture product. This factor mostly considered by project manager as one of the respondent says that;

"Best Suitable architecture for the target product is one of the main part in task allocation process that also needs consideration". Project Manager

Like requirements stability literature also pay little attentions against this factor and only 7% studies considered it as one of the task allocation factor for GSD projects.

The next factors that has more positive response from industrial experts is 'Task dependency' and received 95% positive response. Task dependency affect the task allocation process and also the overall project from many perspectives. Dependent task cannot be allocated to two parallel development site as one task needs the output of other task as an input is one of the main impact regarding task dependency. Task size that receives 87% positive response from industrial experts can also be considered as a combination with task dependency. Any carelessness against these factors delay the project deadline and ultimately increase the overall effort. One of the respondent termed this factor as follows;

"Task dependency put hard constraint in task allocation process and therefore limits the overall benefits of GSD paradigm". Senior Software Engineer

Literature also support the expert's point of view as 44% and 29% studies considered as task allocation factor for GSD projects shown in Table 5.3.

'Resource cost' is one of the important factors from client perspective that received 94% positive response from industrial experts. Explicitly this factor is very important during task allocation process but it also impacted by other factors like task dependency implicitly. Client always try to complete the project within minimum cost by allocating task to low cost site but dependency between tasks limit this benefit. There is also a tradeoff between product quality and resource cost as low cost site may not have required skills for particular task and effect product quality. One experts explain this as;

> *"Tasks are always allocated to low cost site without compromising on product quality. But we have many factors that that limit this benefit.* Software Design Engineer

Literature also support the expert's argument and 47% studies considered it as important factor for task allocation shown in Table 5.3.

'Vendor maturity level' and 'vendor reliability' are the two factors that are more relevant from client perspective and experts give positive response 85% and 77% respectively. These factors are directly related to vendor organization and can be treated from different perspectives. One of the respondent says that;

> *"We always try to divide the work force into recommended size teams in order to avoid the communication and coordination problem".* Software Engineer.

The other perspective that we can link with these two factors is the organization standard and one respondent takes this as;

> *"We implement all the organizations process according to CMMI standards in order to insure better quality".* Business analyst

All these perspective have importance from client perspective while task allocation to any organization as they directly or indirectly impact the product quality and can also delay the project (communication and coordination issues). Literature also support the industrial data and 36% and 21% studies list these factors for task allocation in GSD projects.

The next key factor is 'time zone difference' that received 73% positive response from industrial experts. Mainly time zone difference OR time shift management is concerned with project management but it also affect the task allocation process. Here we have difference in literature it is second ranked task allocation factor, 63% studies considered it one of the important factor while in industry experts ranked it 10^{th} (73% positive response) which depicts that in industry it does not have much influence in task allocation process. The same result can be concluded from last ranked factor from industry experts 'Local Government Rules' received 68% positive response.

Table 5.5 Fisher Exact Test Results

Factors	Vender(n=15)		Client(n=23)		Fisher's Exact Test
	Freq.	%age	Freq.	%age	
Site Technical Expertise	**11**	**73**	**15**	**65**	**0.728**
Time Zone	**9**	**60**	**15**	**65**	**1.000**
Resource Cost	4	27	14	61	0.050
Task Dependency	**6**	**40**	**11**	**48**	**0.744**
Vender Reliability	2	13	12	52	0.020
Task Size	8	53	3	13	0.012
Vendor Maturity Level	1	7	5	30	0.114
Product Architecture	**1**	**7**	**2**	**9**	**1.000**
Requirements Stability	**1**	**7**	**2**	**9**	**1.000**
Local Government Regulations	**2**	**13**	**3**	**13**	**1.000**
Intellectual Property Ownership	**1**	**7**	**0**	**0**	**0.395**

5.3.3. SLR and industrial data Analysis based on Client Vendor Relationship

This section discuss the SLR findings and industrial data analysis based on client vendor relationship. Fisher Exact test of independence is used to measure correlation between client and vendor perspective. Fisher exact test is a special type of Chi square test and it is used when there is to measure the independence between two categories. Fisher exact test results are shown in Table 5.5. Data sheet is attached in appendix A.C.

For industrial data we perform chi square test of independence as there are more than two categories. In chi square test we choose linear by linear association test as this test is more suitable for ordinal data values. Chi square test results are shown in Table 5.6. Data sheet is attached in appendix A.D.

Fisher exact test value is the p-value of the correlation between client and vendor perspective. For this study we use 95% confidence level. From this analysis we can observe that the p-value for factors namely "site technical expertise", "time zone", "task dependency", "Product architecture", "Requirements stability" "local government regulations" and "Intellectual property ownership" is greater than 0.05 which implies that we have association between client and vendor perspective. In subsequent paragraphs each factors is discussed in detail.

Site technical expertise are important for client as well for vendor. This factor is important for client while outsourcing a work unit to any geographical site as each work unit requires particular expertise. With reference to vendor, as vendor needs to maintain required product quality and it can be achieved by allocation task to relevant resource (personal has required expertise). The p-value of literature findings is 0.728 and 0.761 for literature findings, which are greater than 0.05, shown in Table 5.5 and Table 5.6 respectively, and we will accept Null hypothesis and reject alternative hypothesis, hence we conclude that there is association or they is no difference between client and vendor perspective for site technical expertise.

Table 5.6 Chi Square Test Results of Industrial Data

Factors	Occurrence in survey (n=62)										
	Client (n=20)				Vendor (n=42)				Chi-square Test (Linear-by-Linear Association) $\alpha = .05$		
	SA	A	D	SD	SA	A	D	SD	X^2	Df	p-value
Site Technical Expertise	**8**	**12**	**0**	**0**	**16**	**25**	**1**	**0**	**0.092**	**1**	**0.761**
Time Zone	**4**	**14**	**0**	**2**	**10**	**17**	**13**	**2**	**0.544**	**1**	**0.461**
Resource Cost	13	7	0	0	14	24	4	0	6.241	1	0.012
Task Dependency	**12**	**8**	**0**	**0**	**21**	**18**	**3**	**0**	**1.131**	**1**	**0.288**
Vender Reliability	12	8	0	0	20	9	12	1	4.038	1	0.044
Task Size	7	8	5	0	22	18	2	0	4.148	1	0.042
Vendor Maturity Level	10	10	0	0	13	21	8	0	4.382	1	0.036
Product Architecture	**9**	**11**	**0**	**0**	**17**	**22**	**3**	**0**	**0.550**	**1**	**0.458**
Requirements Stability	**12**	**8**	**0**	**0**	**20**	**20**	**1**	**1**	**1.344**	**1**	**0.246**
Local Government Regulations	**4**	**7**	**8**	**1**	**5**	**20**	**16**	**1**	**0.002**	**1**	**0.963**
Intellectual Property Ownership	**7**	**12**	**1**	**0**	**9**	**26**	**6**	**1**	**2.405**	**1**	**0.121**

Time zone difference are also important for both client and vendor. Typically time zone difference enforce asynchronous communication between client and vendor organizations that increase the communication and coordination problems which leads towards delays and effect the software development process (task allocation). Accordingly to minimize the delays and improve task allocation assignments in GSD paradigm client and vendor needs a synchronous mechanism for coordination. Statistical results implies that we will accept null hypothesis and reject alternative hypothesis as p-value of literature findings is 1.000 and 0.461 industrial data, shown in Table 5.5 and Table 5.6 respectively, which are greater than 0.05. Hence we conclude that there is no difference between client and vendor perspective for Time Zone difference as task allocation factor.

Typically interdependency exist between projects tasks that needs to resolve to smooth the work distribution process and maximize productivity. Both needs to make sure that before the start of certain task all previous task should be completed to complete the project on time. Any sort of conflict either on client side or vendor side would cause delays and decrease productivity. Statistical results also support this argument as p-value of literature findings is $0.744>0.05$ and $0.288>0.05$ of industrial experts' opinion, shown in Table 5.5 and Table 5.6 respectively, which implies that we will accept null hypothesis and reject alternative hypothesis. Hence we conclude that there is an association between client and vendor perspective against task dependency factor for task allocation process.

Requirements stability is the measure of the change in requirements when project is in progress. Any change in requirements require the revolution in project planning that increase the overall effort as well as resources and effect the client and vendor equally. This affect can be mitigated by applying proper change management policy in software development process. The p-value of Industrial experts' opinion is 0.246 which is greater than 0.05 shown in Table 5.6, so we will accept the null hypothesis and reject the alternative hypothesis and conclude that there is no difference between client and vendor perspective.

However literature lacks in this domain as there is only 3 (8%) discuss about this concept, so we cannot make any reasonable assumption, but based on current results i.e. p-value is 1.000>0.05, we will accept the null hypothesis and reject the alternative hypothesis means there is no difference between and vendor perspective. Product architecture are also interlinked with requirements stability and literature and industry data both consider it equally important for task allocation process. P-value for expert's opinion is 0.369>0.05 and for literature is 1.000>0.5 but only 3% literature report it as important factor. So we will accept the null hypothesis and reject the alternative hypothesis means there is no difference between and vendor perspective, but as a conclusion these both needs attention in future research.

Local government Regulations equally influence the task allocation process for client and vendor. Labor policies, working hour regulations, payment policies to employees etc. are some sub factors that imposed by local government on local organization. Client and vendor needs to consider all these sub factors while doing task allocation. Statistical analysis of SLR findings and Industrial data also support this concept as p-value is 1.000 and 0.963 which is greater than 0.05, shown in Table 5.5 and Table 5.6 respectively, so, we will accept null hypothesis and reject alternative hypothesis. Hence we conclude that there is no difference between client and vendor perspective against local government regulations factor.

With reference to Resource cost, which is one of the main motivation for client organization behind project outsourcing in GSD paradigm. Clients are mostly interested in low cost development sites and is more concerned instead of vendor organization. Our statistical analysis of industrial experts opinion and literature findings also enforce this concept and it is also worth noting that client studies have more trend towards resource cost than vendor studies, shown in Table 5.5 and Table 5.6. Based on statistical results p-value of literature findings and industrial experts is less than 0.05, so, we accept alternative hypothesis and reject null hypothesis. Hence we conclude that there is no association OR there is a difference between client and vendor perspective for resource cost. However client needs to consider the other factors like site technical expertise and task dependency along with cost.

Organization structure, HRM process, organization employment policies impact the software development process. Collectively all these factors can be termed 'Vendor maturity level' and 'vendor reliability' and client have more concerned regarding factors. These factors can be considered as previous vendor history and positive behavior of these factors increase client trust on vendor and finally helps client while choosing vendors for outsourcing project. For example intellectual property ownership directly related with vendor reliability SLR findings and industrial experts' opinion have more trend towards client perspective which shows client concern for these factors, shown in Table 5.5 and Table 5.6. The p-value of both statistical results is less than 0.05 except vendor maturity level in literature, so, we will accept alternative hypothesis and reject null hypothesis. Hence we conclude that there is a difference between client and vendor perspective about vendor maturity level and vendor reliability but client is more concerned about these factors.

Intellectual property ownership is one of the main factor that insure the success of GSD paradigm. Typically client has to share its ideas with vendor organization in GSD, vendor organization needs to pay attention to secure client intellectual property and it is directly related to task allocation process. Researchers does not show much interest in this perspective as only one study report this as influencing factor for task allocation shown in Table 5.5, so, we cannot make any assumption and it needs attention in future research. However industry expert's data analysis shows that 87% respondents gives positive response. Regarding statistical analysis the p-value of industrial data is 0.121>0.05 shown in Table 5.6 which implies that we will accept null hypothesis and reject alternative hypothesis. Hence we conclude that there is no difference between client and vendor perspective.

5.3.4. Comparison between SLR and industrial data

In this study we collect data of task allocation factors in GSD environment from two data sources, SLR and Questionnaire survey and this section presents a comparative analysis of the challenges identified by the these data sources. In order to quantify the significance of the similarity in the factors identified using the SLR and the questionnaire-based empirical study, we performed a T test of independence as data came from two different data sources [98].

In order to analyze similarities between two data sets, two hypothesis are designed. One is null and other is alternative hypothesis. In t-test two-step process is used to accept or reject the hypothesis. First check the p-value of Levene's Test for Equality of Variances, if this is less than 0.05 then we will consider the option 'equal variance not assumed' otherwise we will consider ‘equal variance assumed’. Hypothesis acceptance OR rejection is based on P-Value and confidence level is set at 95%.

H1 (Null Hypothesis): There is an association between two data sets (SLR, Questionnaire survey) OR both are same.

H2 (Alternative Hypothesis): There is no association between two data sets (SLR, Questionnaire survey) OR both are different.

Table 5.7 Group Statistics

	Type	N	Mean	Std. Deviation	Std. Error Mean
Factor	SLR	11	30.6364	22.94024	6.91674
	Survey	11	41.2727	10.20873	3.07805

Table 5.7 shows the descriptive statistics of two data sets used for this study. Table 5.8 shows independent sample T test results. The columns labeled "Levene's Test for Equality of Variances" tell us whether an assumption of the t-test has been met. The t-test assumes that the variability of each group is approximately equal. If that assumption isn't met, then a special form of the t-test should be used.

Table 5.8 Independent Samples T Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Factor	Equal variances assumed	8.67	.008	-1.405	20	.175	-10.63636	7.57071	-26.428	5.15587
	Equal variances not assumed			-1.405	13.81	.182	-10.63636	7.57071	-26.894	5.62204

In this study the p-value for levene's test is 0.008<0.05 so we will consider option 'equal variance not assumed. Now in this option we will check p-value to accept or reject the hypothesis. The p-value for this option is 0.182>0.05, so, we will accept the null hypothesis and reject the alternative hypothesis, means that there is an association between these two datasets OR both are same. Hence we will conclude that;

"A t test failed to reveal a statistically reliable difference between the SLR factor data that has (M = 30.63, s = 22.94) and the Questionnaire survey data that has (M = 41.27, s = 10.20), t (13.81) = 1.405, p = .182, α = .05."[98]

CHAPTER 6

TASK ALLOCATION MODEL

6.1. Introduction

This chapter explain the task allocation model for GSD projects. In this model we consider two factors "site technical expertise" and "task dependency" for task allocation. These factors are based on literature findings and industrial expert's data. We run task allocation on one real time project (Life Cycle Assessment (LCA) Tool) as well as one problem instance from project scheduling problem repository. Section 6.2 discuss the task allocation basis and section **Error! Reference source not found.** describe the results of case studies.

6.2. Task Allocation Model Description

Global Software development life cycle differs in many perspectives from in-house software development. Task allocation is one of the main activity during initial phase of development life cycle. In GSD different geographical sites are involved in software development and tasks needs to allocate to these different geographical sites. This section explains the task allocation model for GSD environment that that make use of tasks details (project schedule), GSD site characteristics and allocate tasks to different GSD sites based on task allocation objective function.

The task allocation process for GSD projects are shown in Figure 1, consist of four phases namely, 'project schedule', 'GSD site characteristics', 'task allocation objective function' and 'portfolio of optimal solutions'. The first phase project schedule includes task title, duration in terms of days, man hours required in term of days, and expertise for that particular task (past experience, programming skills). The second phase use GSD site characteristics includes man hours available on particular site per day and available expertise (past experience, programming skills). The third phase describe the task allocation objective function, which is a combination of goals (skill set matching, work load matching and resource leveling), and use Genetic algorithm to generate optimal solutions. As a result a portfolio of optimal solutions for each geographical site is identified. The four logical phases are discuss in details in subsequent sections.

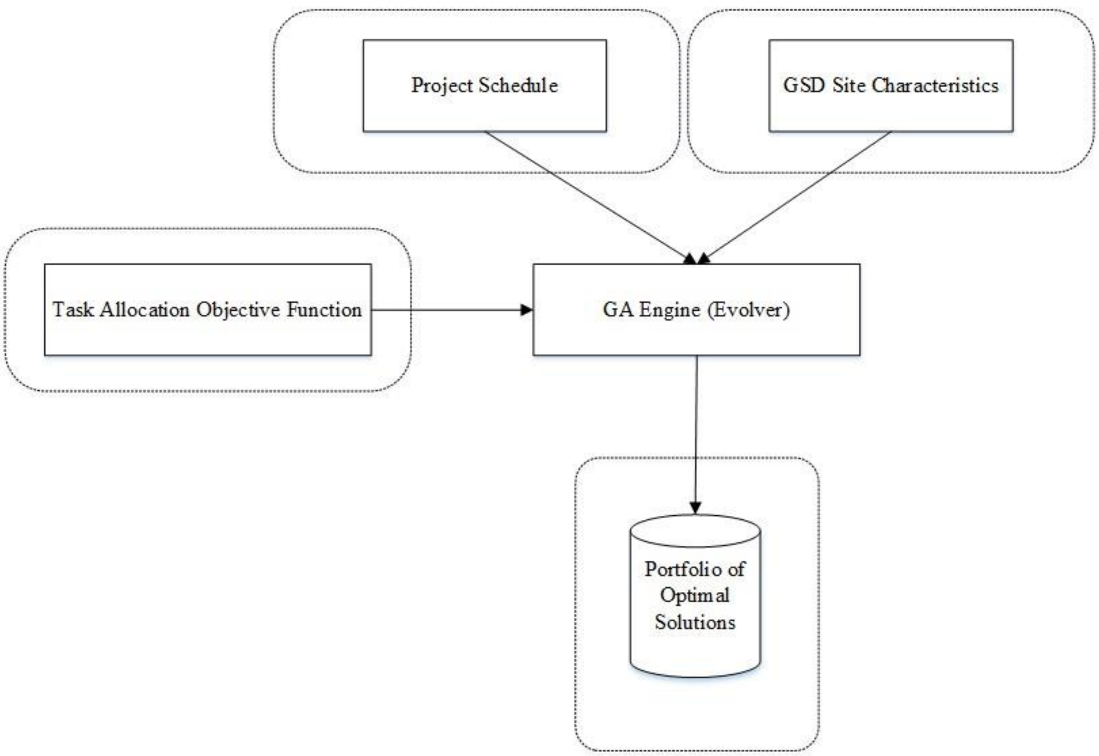

Figure 6.1 Task Allocation Model Overview

6.2.1. Project Schedule

This section describe the first phase of task allocation process 'project schedule'. Project schedule includes tasks precedence, duration in number of days, required man hour requires per day, and required expertise (past experience and programming skills). Five level scale is used to represent expertise, one means that lowest level and five means the highest level of expertise available on particular site. Based on the precedence and duration details for the tasks, critical path method (CPM) or program evaluation and review technique (PERT) is applied to get the project schedule. CPM is utilized for the projects which are very common in the domain and where a single duration value of the tasks can be provided [99]. On the other hand PERT is used for the project tasks` where there is uncertainty in the duration of the tasks[100]. For such tasks three sorts of durations are described namely minimum possible duration, maximum possible duration and most likely duration. Using triangular distribution, the calculations are made for the schedule. Once the schedule is made we have earliest start, latest start, earliest finish and latest finish dates for a certain project tasks. Based on these dates a float can be calculated. By float, it means the maximum duration for which a task can be delayed without delaying the end date of the project. This float allows the flexible dates for the start of a task without delaying the deadline of the project. It is important to mention that float is only available for non-critical activities. By critical activities, it mean the activities which play critical role in determination of the project`s duration. So in the process, the project manager finds the room to adjust the non-critical activities in order to satisfy a certain tasks` allocation criteria. It is also important to mention that during all of the above process described above, the precedence relationship among the activities is preserved to maintain the logic of the activities` network.

6.2.2. GSD Sites Characteristics

This section discus the second phase of task allocation model' GSD site characteristics'. In global software development typically different geographically development sites are involved and task needs to allocate these different sites. Each site has certain characteristics that plays important role during task allocation process. In our process model, man hour available per day, particular available expertise include past experience, programming skills are some properties of sites that are used in task allocation process. Five level scale is used to represent expertise, one means that lowest level and five means the highest level of expertise available on particular site.

6.2.3. Task Allocation Objective Function

This section explains the third phase of task allocation process 'task allocation objective function'. In this study task allocation function comprises a set of sub functions that we apply individually and collectively in successive stages to achieve different objectives. Two sub functions are 'skill set matching' and 'work load matching' and also we apply another constraint 'resource leveling' to improve these sub functions. Subsequent sub sections describe each sub function in detail.

(i). Skill Set Matching

Typically a software tasks require certain set of expertise i.e. programming skills, platform skills etc. in our study we divide the required expertise for each into two categories namely ‘past experience’ and ‘programming skills’, past experience can be considered in terms of environment OR platform experience, and domain experience etc. As each task require certain type of past experience and programming skills that should be available on particular site to which it will be assigned. On the other hand we also have certain expertise available on each GSD site. The main purpose of this sub function is to enhance the productivity and improve the quality of the product by minimize the difference between required skills for particular task and available skills on particular site.

Skill matching criteria tries to match the required skills-set for a task to a site where similar skills-set is available. Mathematically this is achieved by minimizing the following expression

$$\sum_{i=0}^{j}\left(\sum_{m=1}^{n}\frac{|SK_{im}-S_{im}|}{4\ x\ im}\right) \qquad (1)$$

Where j are the tasks for a particular project, SK is the skills-set required for a certain task and S is the skills-set available at a certain GSD site and m are the number of skills included in the task allocation process. By using this formula we make 0-1 scale for set of expertise i.e. in this study we have two type of expertise so each expertise has value from 0 to 0.05. 0 means required and available skills are exactly matched and 0.05 means there is no match between requirement and availability of expertise.

Although the skills set matching is prioritized for this scenario but some other constraints are fixed for a real life scenario. These constraints include

- The overtime allowed for a particular day is limited to 140% of available man-hours on a specific day on a particular site. Further the total overtime at a specific time is limited to 120% of normal man-hours available throughout the span of the project. This assumption is made by considering the fact in mind that the extended overtime on the project may result in decrease in the productivity of the resources.
- Work was distributed to the sites while keeping in mind that a specific minimum work load was transferred to all of the working sites. This minimum working load in terms of man-hours was obtained by dividing the total man-hours of the project by number of global development sites and a variable. This variable could be changed to vary the minimum work load criteria. For our scenario, we kept it to be "2" so that minimum work load was 50% of the work load on any site if the work is distributed evenly to the all of the sites.

(ii). Work Load Matching

In our study along with required expertise for task, there is another constraint 'required man hours per day' that needs to be consider during task allocation process. In first case we consider only product quality perspective and try to minimize the difference between required skills for particular task and available skills on particular site. While improving quality by satisfies the skill set requirement we may impose extra load than available on particular site that ultimately degrade the overall performance. In order to address this deficiency we try to minimize the difference between available man hours per day and required man hours per day on particular site along with skill set matching. In this scenario we combine sub function 'skill set matching' with 'work load matching' and try to optimize them. The purpose of the second scenario is to achieve quality while considering the project management perspective simultaneously. The quality objective is same as was in the first scenario that is try to match the skills set required to the skills set available at different sites. For project management perspective, the model tries to matches man-hours required for a certain task to man-hours available at different sites. In this case we have more than one goal and for such cases goal programming plays a critical role [100].

Work load matching criteria tries to match the tasks work load to the sites where similar man-hour resources are available. In this way the model tries to reduce the overtime of the resources for the project tasks throughout the span of the project. Mathematically it is achieved by minimizing the following expression

$$\sum_{1}^{s}\sum_{1}^{\mathrm{d}} \mathrm{stdev}\beta \tag{2}$$

Where, β =Standard deviation between work load assigned and working capacity on that day at that particular site to which work is assigned in term of man-hours, d are the number of days for the project and s are number of geographical sites available for the software development. This sub function can be considered as multi objective function.

In this case we apply same two constraints as in first case but additionally we apply one another constraint which is as follows.

- The total overtime at a specific time is limited to 120% of normal man-hours available throughout the span of the project.

By applying these constraints got required results but the use of resources is not smooth throughout the project time span. So in order to overcome this problem we apply one special constraint ‘resource leveling’. The subsequent discuss this constraint in detail.

The other constraint

Resource Leveling: As in work load matching we try to minimize the difference between required man hours for task and available man hours on particular site but we did not pay attention to work load balancing throughout the whole project life span. This type of working plan is mostly not acceptable for project managers as it demands more work forces in the start and then less at the middle and end. In order to overcome this deficiency we apply another criteria which we used in our model is resource leveling. The concept behind resource leveling is to try to balance the work load equally throughout the span of the project. This helps the project to overcome the requirement of overstaffing during the peak load days and assign more tasks when the resources have free time during their job duties. In other words work load distribution is not stable through the project time span. We can overcome on this deficiency by using the slack i.e. different between early start (how much early specific task can started) and late start (how much we can late the start of particular task without affecting the schedule) [99]. This adjustment can be done for only non-critical path tasks. The purpose constraint is to improve the task allocation process purely from the project management perspective i.e. to allocate the tasks in such a manner that work load is divided almost equally throughout the span of the project.

Mathematically this is achieved by minimizing the following expression.

$$\text{stdev}\sum_{1}^{d}\mu \qquad (3)$$

Where. μ= Work assigned on a particular day in term of man-hours across all of the sites.

The recipe used for task allocation can either use one of the allocation priority criteria or may use multiple of them at a same time. This is achieved by goal programming technique as already mentioned.

6.2.4. Portfolio of Optimal Solutions

This section describe the output of our model 'portfolio of optimal solutions'. Once we have the project schedule, GSD site characteristics and task allocation objective function, we will apply genetic algorithm on this set of inputs and generate a set of optimal solutions for task allocation to geographical sites. The project manager can use any solution from given set of solutions based on his own experience. The project manager can change any of the sub functions or put extra constraints based on his interest function.

6.3. Application of Task Allocation Model

This section presents the application of proposed task allocation model in GSD environment. Section 6.3.1 discuss the case studies that we use to validate our model and section 6.3.2 discuss the results of these case studies.

6.3.1. Experimental Design

In this study the proposed model is applied on two case studies, one case study is real time project Life Cycle Assessment (LCA) Tool [101]. The second case study is randomly generated instance from Project Scheduling Problems Library (PSPLib) [102]. Project

Scheduling Problems Library that researcher used to implement and validate scheduling problems solutions.

In this study we only consider development phase of software development life cycle. Life Cycle Assessment Tool (LCA) case study is consist of 34 development tasks and each task has set of requirements like duration, required man hour per day and required expertise (required past experience and required programming skills).

Table 6.1 presents the LCA tasks list, planned duration in days, required programming and toll skills and their dependencies, Figure 6.2 represents LCA CPM network.

Table 6.1 LCA Tasks

Task ID	Activities	Duration	Predecessors				Man- Hour/day	Past Experience Required for task	Programming skills Required for task
1	Define Actor	2					5	3	5
2	Product scope and goal	3	1				5	2	2
3	Create Database	11	2				4	5	1
4	Import Csv OR SimaPro files	6	2				6	4	5
5	Import .txt OR Excel files	8	2				5	3	4
6	Solve Conversion Problems	5	3	4	5		4	4	2
7	Create Flows	7	6				3	5	1
8	Modify Flows	5	6				5	2	5
9	Create Process	4	7	8			6	3	4
10	Integrate Business logic	10	9				4	3	3
11	Implement GUI production phase	15	2				5	4	5

12	Integrate GUI production phase and Business Logic	7	10	11			5	5	5
13	Import LCA Methods	3					4	1	1
14	Sequential Inventory calculation	10	13				6	2	1
15	Uncertainty Calculation	11	13				3	3	4
16	Characterization	6	14	15			2	5	5
17	Damage Assessment	8	14	15			7	4	2
18	Normalization	7	14	15			5	3	3
19	Characterization	8	14	15			6	4	2
20	Group Analysis	10	16	17	18	19	4	5	5
21	Standard Analysis	12	16	17	18	19	5	2	5
22	Graphical Analysis	13	20	21			4	3	5
23	Spread Sheets analysis	14	20	21			6	4	4
24	Export Result in .csv OR SimaPro Form	7	22	23			6	5	3
25	Export Results in .txt OR Excel form	6	22	23			3	3	2
26	Integrate Analysis Phase	15	24	25			1	5	1
27	Implement analysis phase GUI	17	13				8	4	2
28	Integrate GUI with Business Logic	16	26	27			9	3	2
29	Import product	6					15	2	3
30	Compare two product systems	9	29				2	1	5
31	Modify Product System	10	29				5	4	4
32	Create project from different product systems	8	30	31			6	5	2
33	Integrate Complete system	22	12	28	32		6	5	5
34	Deploy system in testing environment	12	33				6	5	3

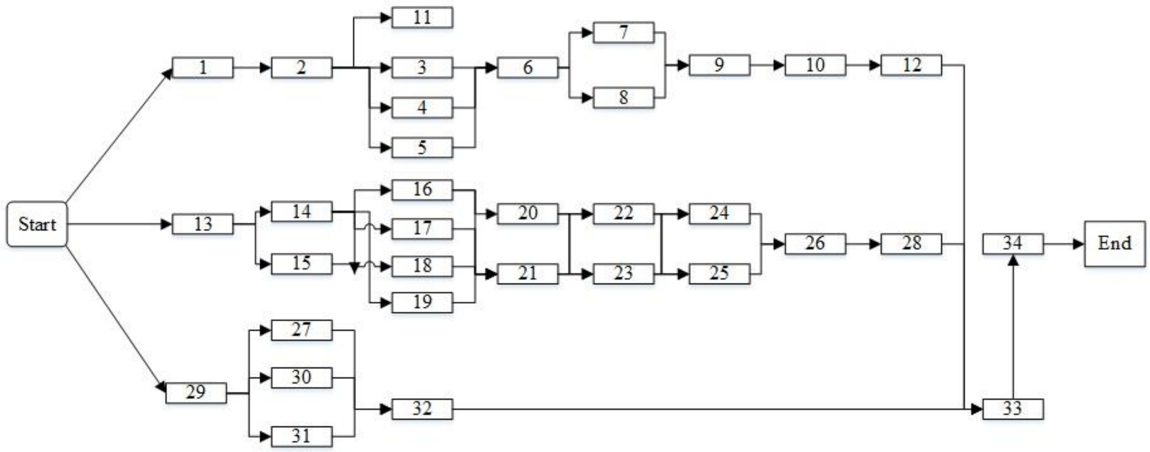

Figure 6.2 LCA Activity Network

Table 6.2 Shows the GSD sites characteristics includes Man-hours available per day, Tool Expertise and programming skills availability. We use six geographically different GSD sites for this project.

Table 6.2 GSD Site Characteristics (Case Study-1)

Geographical Sites	A	B	C	D	E	F
Resources Available (man-hours/day)	6	4	5	3	6	7
Past Experience	4	5	2	5	5	3
Programming Skills	4	3	5	5	2	4

In order to validate the efficiency and extendibility of our model we take one case study form project Scheduling Problem Library (PSPLib)[102]. Many researchers use this well-organized library to validate his work. We take one data instance of 60 activities, across each activity or task there is certain set of requirements including duration, required man hours per day and required expertise, and total duration is 77 days. Case study 2 task details are shown in Table 6.3 . Figure 6.3 shows the activity network of that data instance.

Table 6.3 Case Study 2 Tasks Details

Task ID	Duration	Predecessors				Man-Hour/day	Past Experience Required for task	Programming skills Required for task
1	8	1				10	3	5
2	1	1				1	2	2
3	10	1				9	5	1
4	6	2				4	4	5
5	5	5				1	3	4
6	8	3				10	4	2
7	9	4				6	5	1
8	1	8				8	2	5
9	9	2				6	3	4
10	8	10				3	3	3
11	3	4				7	4	5
12	6	9				8	5	5
13	2	3				1	1	1
14	5	2				9	2	1
15	1	4				6	3	4
16	3	6				2	5	5
17	10	13	14			2	4	2
18	9	14				7	3	3
19	1	8				5	4	2
20	3	12				8	5	5
21	6	5	15			4	2	5
22	3	7				4	3	5
23	3	5				1	4	4
24	7	15				5	5	3
25	6	11				9	3	2
26	10	12	20			7	5	1
27	9	18	19			3	4	2
28	8	3				3	3	2
29	4	27				7	2	3

30	3	20	22			6	1	5
31	3	17				4	4	4
32	6	18				7	5	2
33	1	14				4	5	5
34	9	9				1	2	5
35	9	35				9	5	2
36	1	11				7	4	1
37	2	6				5	3	5
38	4	21	33	35		1	4	4
39	9	8	25			5	5	2
40	10	29				1	2	1
41	8	36				9	3	5
42	4	17	31			6	3	4
43	3	11	34			1	4	3
44	6	10				9	5	5
45	6	41				7	1	5
46	7	28				4	2	1
47	3	24	42			8	3	1
48	2	26				2	5	4
49	10	38	44			7	4	5
50	4	23	27	43		5	3	2
51	2	47				2	4	3
52	1	39	40	45		1	5	2
53	4	26	52			6	2	5
54	10	16	48	54		7	3	5
55	8	30	46	53		3	4	5
56	6	29	32	55		4	5	4
57	10	16	37	43		9	3	3
58	3	15	49	57	58	7	5	2
59	10	38	51			3	4	1
60	1	28	50	56		1	3	3

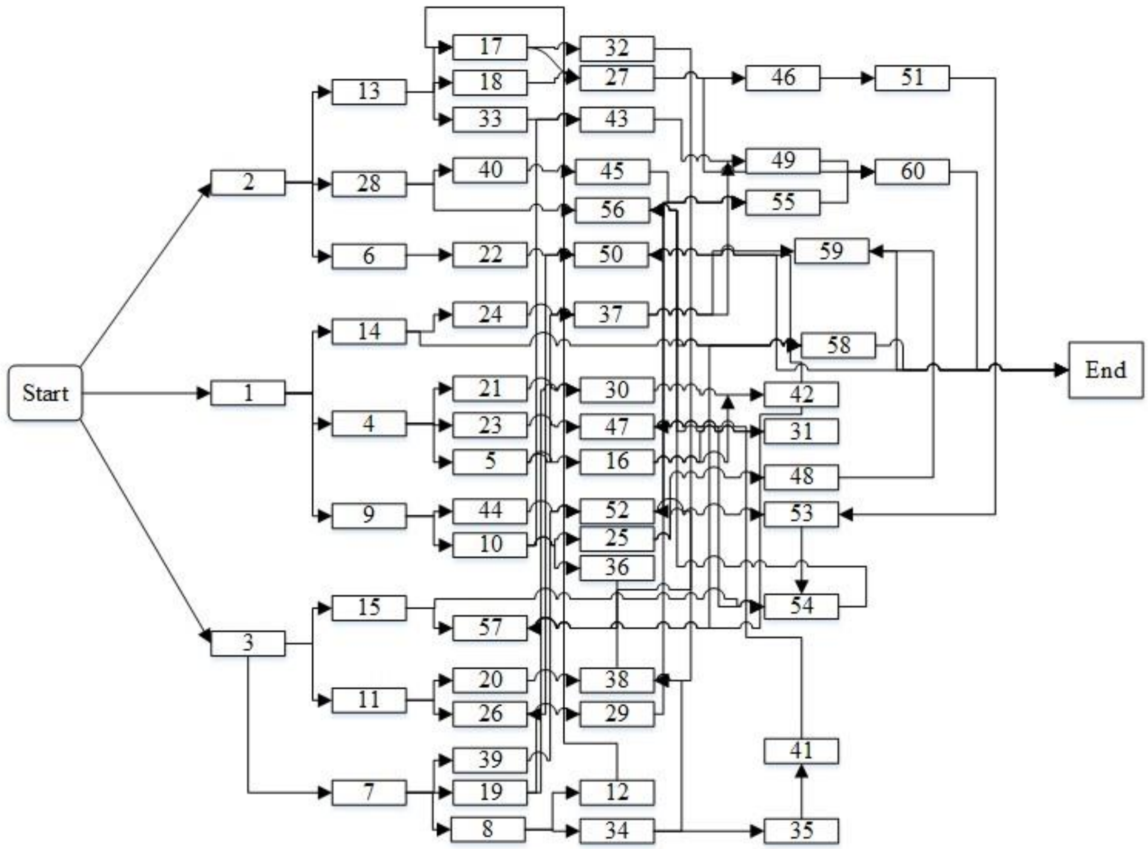

Figure 6.3 PSP Instance Activity Network

Table 6.4 Shows the GSD sites characteristics includes Man-hours available per day, Tool Expertise and programming skills availability. We use four geographically different GSD site for this project.

Table 6.4 GSD Site Characteristics (Case Study-II)

Geographical Sites	**A**	**B**	**C**	**D**
Resources Available (man-hours/day)	10	8	9	7
tool expertise	4	5	2	5
programming skills	4	3	5	5

6.3.2. Results and Discussion

This section discuss the results of case studies that we use to validate task allocation model. In each case study we apply two cases that we discuss in previous section with specified constraints. All possible scenarios are discussed in subsequent section in detail.

(i). Skill Set Matching

The main purpose of this sub function is to enhance the productivity and improve the quality of the product by minimize the difference between required skills for particular task and available skills on particular site.

With reference to case study one, Figure 6.4 shows the initial solutions of skill set matching. On x-axis are the development tasks while on the y-axis, there is a scale showing skills set matching. Since we had only two skills in this scenario so each skill i.e. past experience and programming skill matching had 50% weightage each. The scale is from 0 to 0.5, 0 indicating exact match between the required skills set and skills set available on a site where the task was assigned while 0.5 indicates that there is no matching for the skills set. The initial solution was obtained by solving the constraints (Dependency and overtime) imposed on the case as described above.

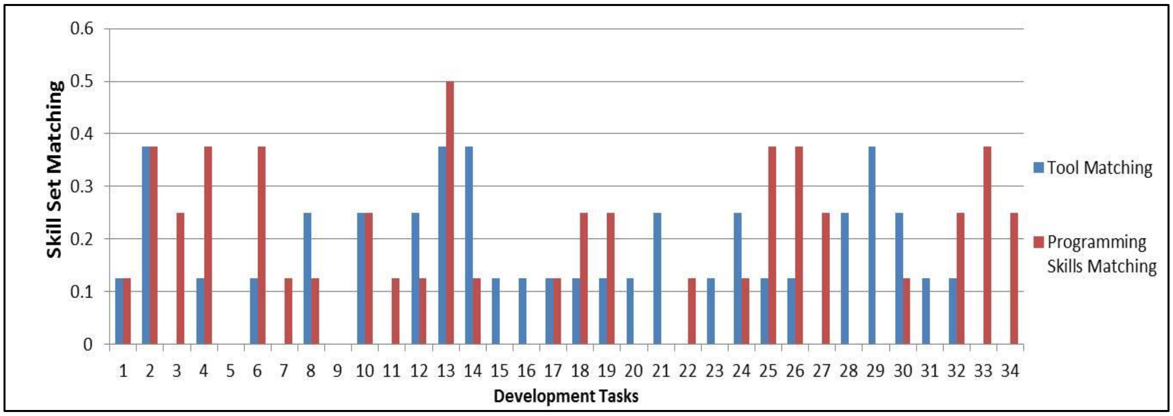

Figure 6.4 Initial individual Skills Matching-1

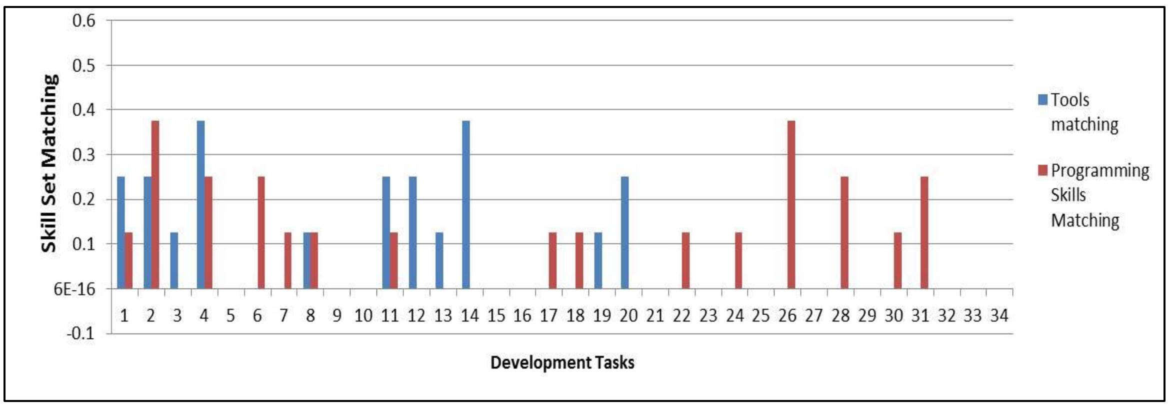

Figure 6.5 optimized individual Skills Matching-1

After solving all constraints we apply optimization function on initial solution. Figure 6.5 shows the skills set matching after running the optimization. The difference is quite obvious in these figures. Skills set is exactly matched (available skills on particular site and required skills for that task are equal) for tasks 10, 15, ,16, 21,23,25,27,29,32,33,ande 34 and for Many of the other task difference between the requirement and availability was minimized.

With reference to case study two, Figure 6.6 shows the initial solution for skill set matching. On x-axis are the development tasks while on the y-axis, there is a scale showing skills set matching. Since we had only two skills in this scenario so each skill i.e. past experience and programming skill matching had 50% weightage each. The scale is from 0 to 0.5, 0 indicating exact match between the required skills set and skills set available on a site where the task was assigned while 0.5 indicates that there is no matching for the skills set. The initial solution was obtained by solving the constraints (Dependency and overtime) imposed on the case as described above.

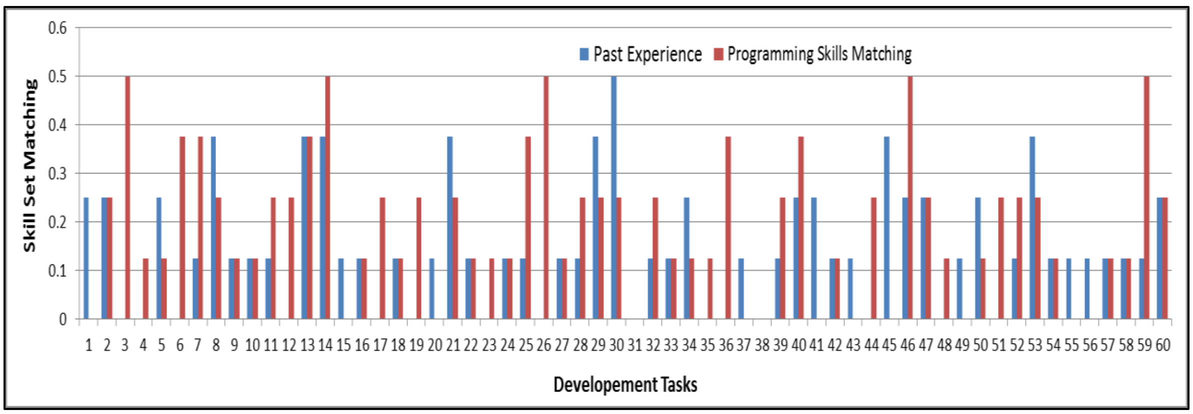

Figure 6.6 Initial Individual Skills Matching

After solving all constraints we apply optimization function on initial solution. Figure 6.7 shows the skills set matching after running the optimization.

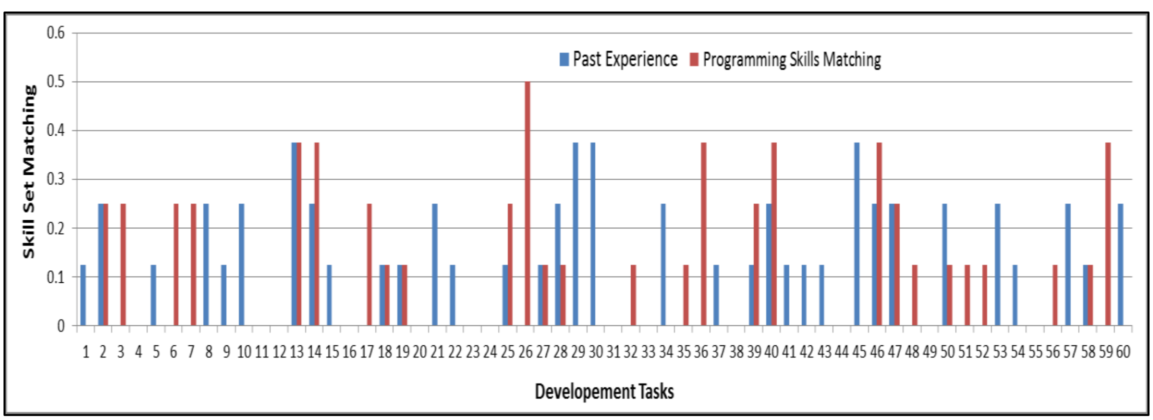

Figure 6.7 Optimized Individual Skills Matching

Skills set is exactly matched (available skills on particular site and required skills for that task are equal) for tasks 4,11,12,16,23,24,31,38,44,49,55, and 56 and for tasks 1,5,9,10,11,15,21,22,29,30,34,37,41,42,43,45,53,54,57 and 60 programming skills are exactly matched , 15, ,16, 21,23,25,27,29,32,33,ande 34 and for many of the other task difference between the requirement and availability was minimized.

Figure 6.8 shows the combined results for the complete skill set (past experience and programming skills) of case study one and matching scale is 0 to 1 as we are analyzing complete skill set. 0 means exact match between requirement and availability and 1 means no match. We can observe that for each task there is improvement in skill set matching after optimization except 2 tasks 1 and 4, and also the big improvement can be seen at task 13 where before optimization the value is almost 0.9 (minimum match between required skills and available skills on particular site) while after optimization we have a value 0.1 (means little bit difference between required skills and available skills). The difference in skills set matching is clear by distinguishing the blue bars for “before optimization” and red bars for “after optimization” results.

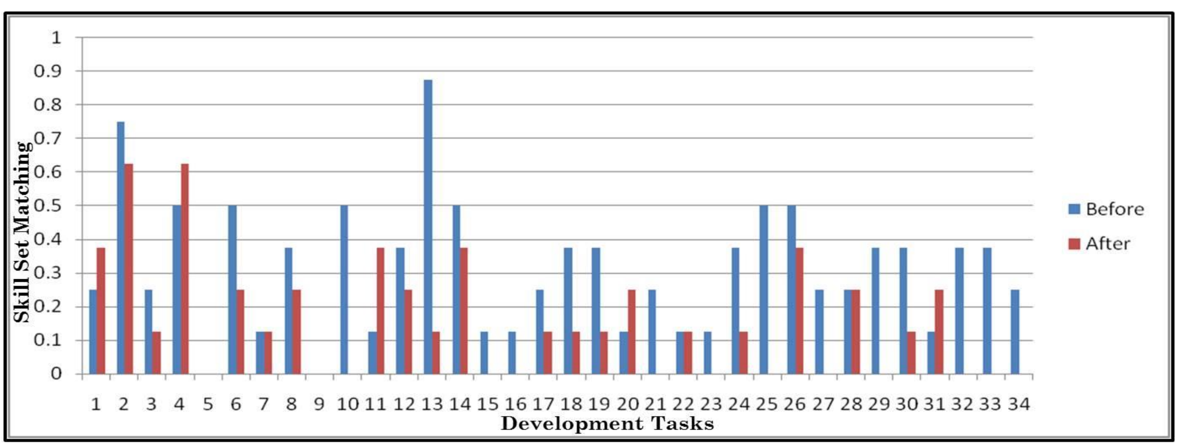

Figure 6.8 Skill Set Matching-1

Figure 6.8 shows the combined results for the complete skill set (past experience and programming skills) of case study two. We can observe that for tasks 4,11,12,16,20,23,24,33,44,49,and 55 skill set is exactly matched after optimization and not even in one tasks we have negative effect of optimization.

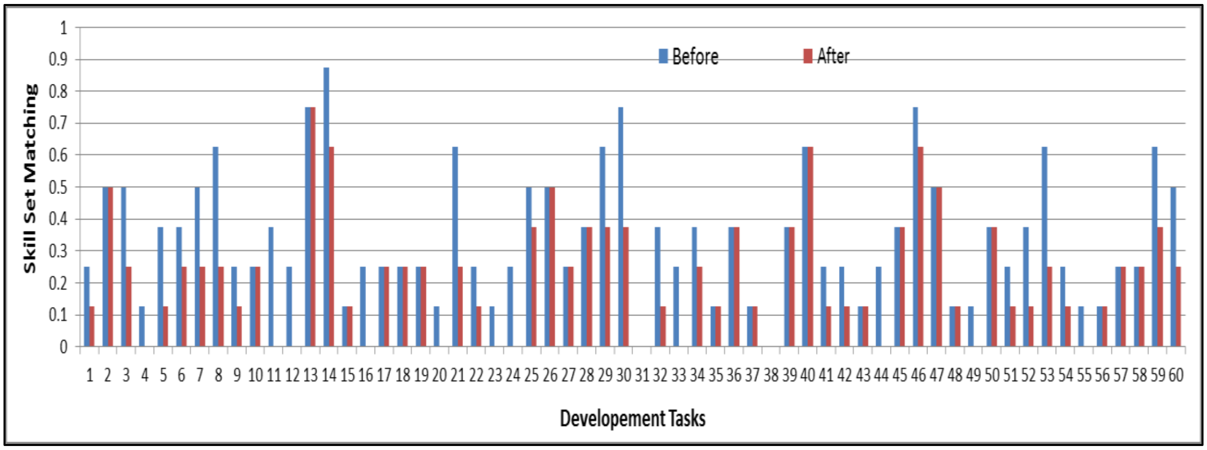

Figure 6.9 Skill Set Matching-2

(ii).Work load and skill set matching

In first case we consider only quality perspective and try to minimize the difference between required skills for particular task and available skills on particular site. While improving quality by satisfies the skill set requirement we may impose extra load on particular site than available on particular site that ultimately degrade the overall performance. In order to address this deficiency we try to minimize the difference between available man hours per day and required man hours per day on particular site along with skill set matching. The purpose of the second scenario is to achieve quality while considering the project management perspective simultaneously. The quality objective is same as was in the first scenario that is try to match the skills set required to the skills set available at different sites. For project management perspective, the model tries to matches man-hours required for a certain task to man-hours available at different sites. For such a scenario, more than one goal was to be achieved.

With reference to case study one, Figure 6.10 shows the initial work load matching. X-axis shows the span of the project in number of days whereas y-axis shows cumulative work load deviation across all sites between work assigned and man-hours capability of a site where task was assigned.

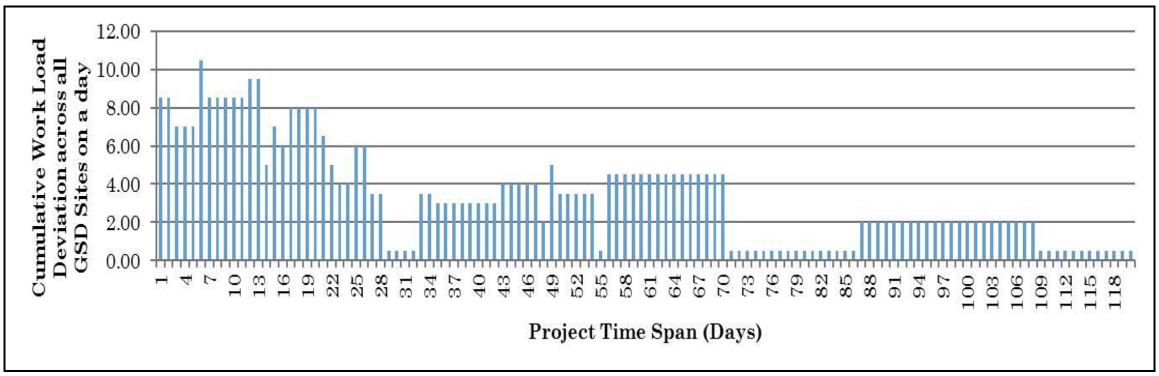

Figure 6.10 Initial Work Load Matching-1

Figure 6.11 shows the results after optimization for work load matching. We can observe that work load deviation is 8 for almost 28 days and 4 for days 34 to 70 before optimization. On the other hand work load deviation is almost 4 for first 25 days and then is near .5 and then almost 2.3 for 55 to 70 days. We also have exact match from 70 to 87 working days. The difference is quite obvious from the two figures for work load matching.

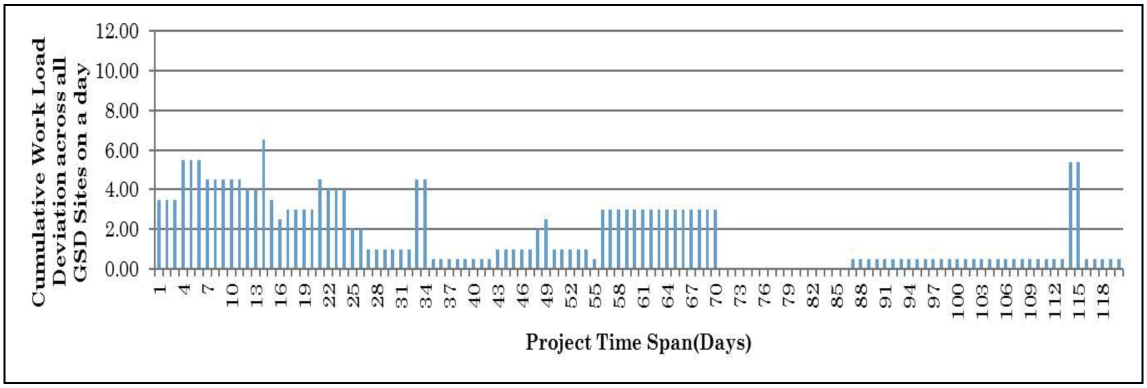

Figure 6.11 Optimal Work Load Matching-1

With reference to case study two, Figure 6.12 shows the initial work load matching. X-axis shows the span of the project in number of days whereas y-axis shows cumulative work load deviation across all sites between work assigned and man-hours capability of a site where task was assigned.

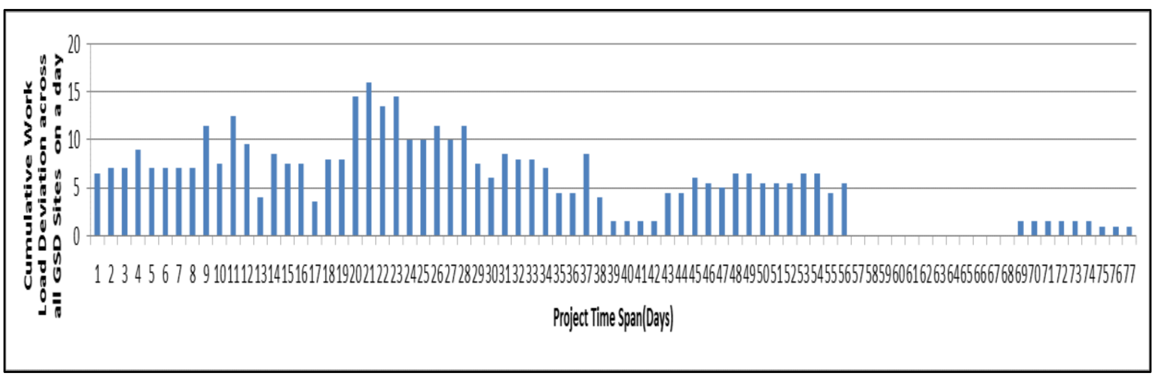

Figure 6.12 Initial Work Load Matching-2

We can observe that cumulative work load deviation is above 5 for almost from almost first 38 days and also between 43 to 56 days before optimization shown in Figure 6.12. On the other hand cumulative work load deviation is above 5 for only 13 days (15,16,17,18,19,20,44,45,46,47,60,61 and 62) for first 25 days and for remaining days the average cumulative work load deviation is almost 3 hours as shown in Figure 6.13. The difference is quite obvious from the two figures for work load matching.

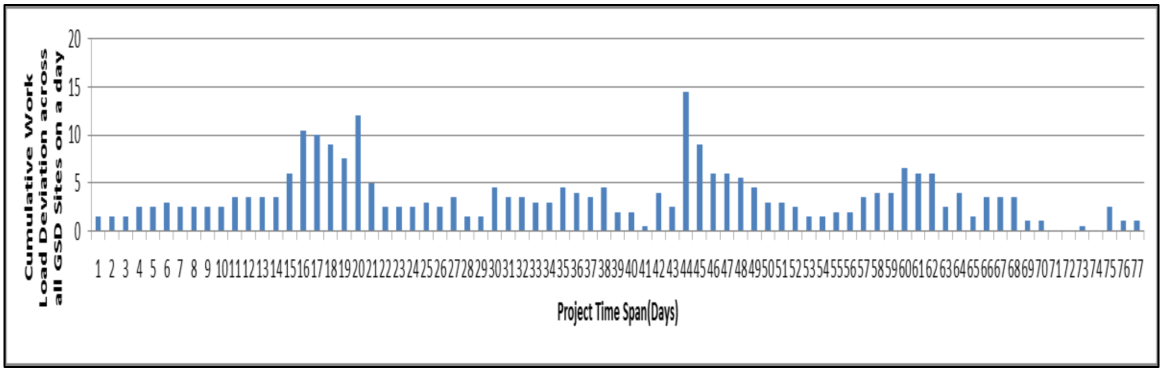

Figure 6.13 Optimal Work load Matching-2

As discussed earlier, for this case we tried to optimized work load matching and skills matching at the same time. When we try to optimize more than one function, the improvement process may not be as clear as compared to the scenario where we had only one function. That is why in this case the skills-set matching is not as obvious as it was in the previous case but still we got improvement in matching for most of the task except some tasks.

Figure 6.14 shows the skills-set matching for case study one. We can observe that for tasks 18, 21, 23, 24, 25 and 33, the skill set is completely matched and for other tasks there is considerable improvement. But for tasks 1,4,7,9, 11, 12, 15 20, 22, 31 and 32 optimization have negative effect i.e. the difference between required skills and available skill has increased.

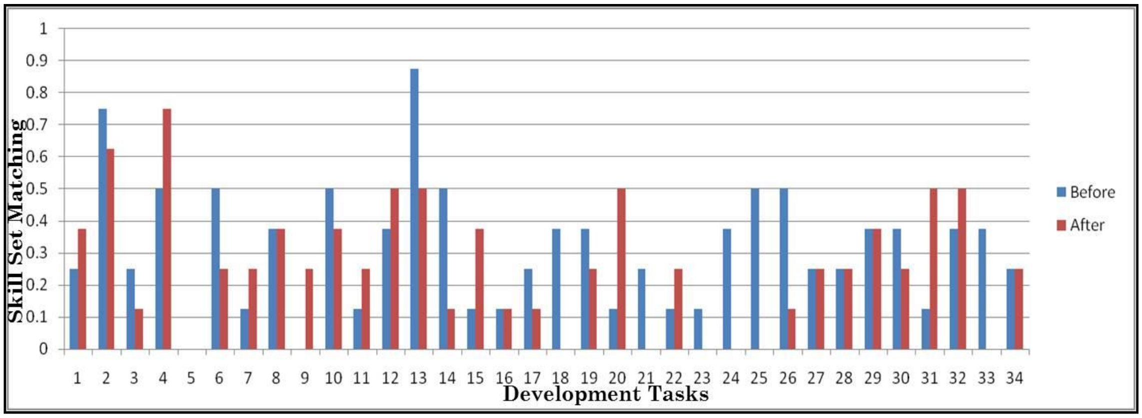

Figure 6.14 Optimal Skill Set Matching-1

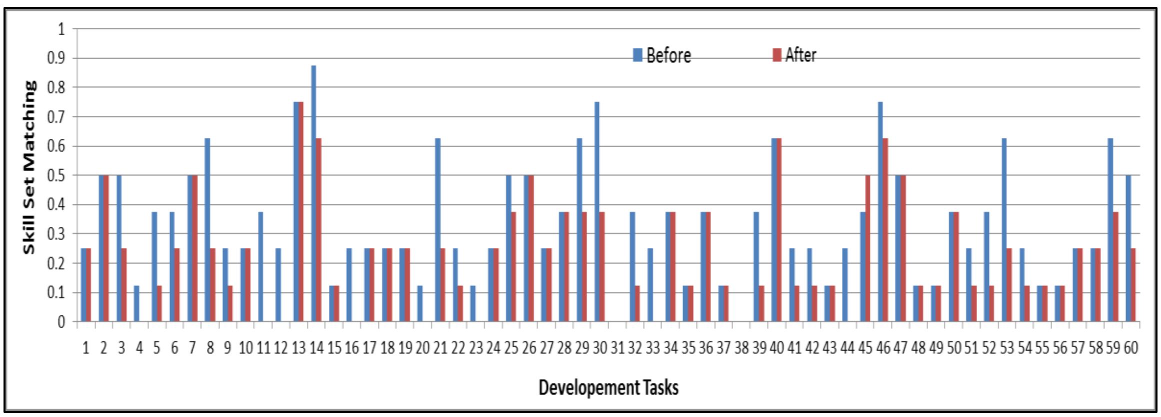

Figure 6.15 Optimal Skill Set Matching-2

Figure 6.15 shows the skills-set matching for case study two. We can observe that for tasks 4, 11, 12, 20, 23, 33 and 44 the skill set is completely matched and for other tasks there is considerable improvement. But for only one task 45 the optimization have negative effect i.e. the difference between required skills and available skill has increased.

(iii). Resource Leveling

The concept behind resource leveling is to try to balance the work load equally throughout the span of the project. This helps the project to overcome the requirement of overstaffing during the peak load days and assign more tasks when the resources have free time during their job duties. In other words work load distribution is not stable through the project time span. We can overcome on this deficiency by using the slack i.e. different between early start (how much early specific task can started) and late start (how much we can late the start of particular task without affecting the schedule)

With reference to case study one, Figure 6.16 shows the work load distribution with respect to project time span before the optimization process. X-axis describes the time span whereas y-axis is showing cumulative work load across the entire sites in term of man-hours for a day.

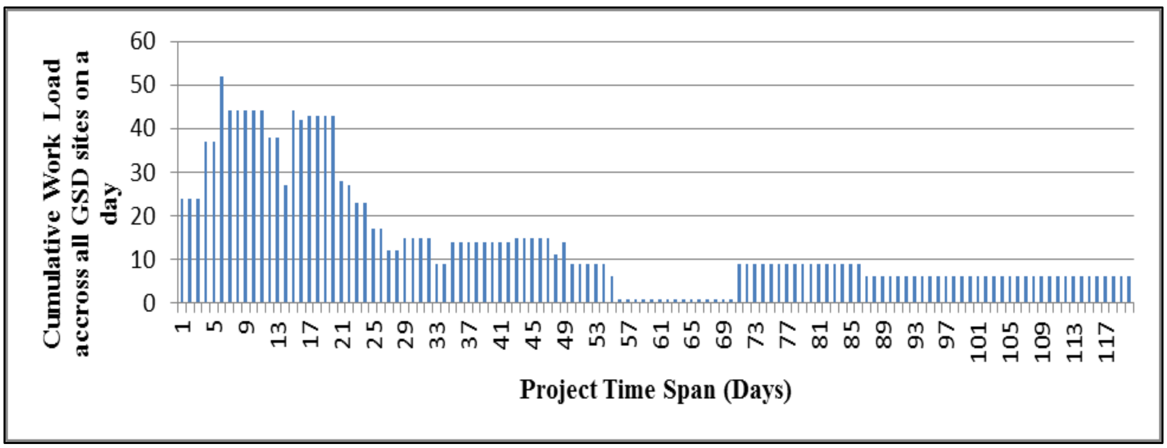

Figure 6.16 Initial Work Load Distribution-1

We can observe that before optimization almost for first 21 days total work load is above 30, from 26 to 53 the total work load on all sites on a single day is about 12, from 57 to 71 day the cumulative work load is negligible and then for remaining days the total work load across all GSD sites is near to 10 man hours. By considering this distribution the project manager cannot make any assumption for his work force as it leads project time span towards un-stable state.

Figure 6.17 shows the optimal work load matching as well as resource leveling i.e. try to distribute the work load equally across whole project time span. We can observe that after optimization the work load is equally distributed to some extent. The cumulative work load across all GSD sites on a particular day is near 10 man hours for most of the days in project time span. Cumulative work load for first 21 days in almost 25 except on day as compared to 40 that we have before optimization. Another major improvement in work load distribution is that now we have no particular days without any workload.

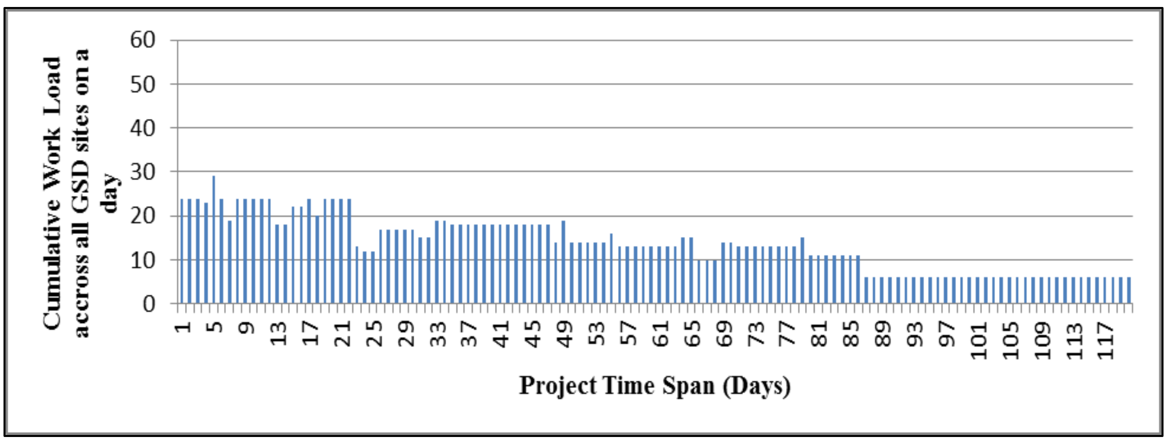

Figure 6.17 Optimal Work Load Distribution (Resource Leveling)-1

With reference to case study two, Figure 6.18 shows the work load distribution with respect to project time span before the optimization process. X-axis describes the time span whereas y-axis is showing cumulative work load across the entire sites in term of man-hours for a day. We can observe that before optimization almost for first 37 days total work load is above 30, from 38 to 50, the total work load on all sites on a single day is between 10 and 20 and for reaming days the total work load across all GSD sites is near to 10 man hours.

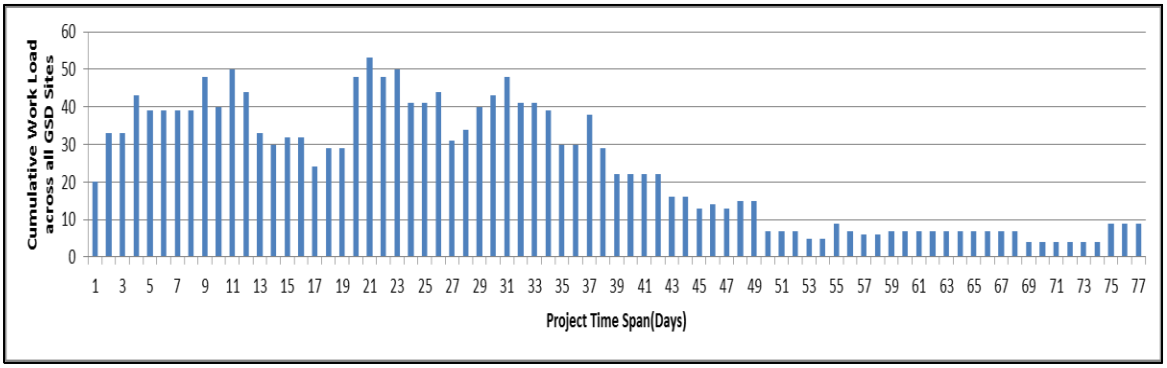

Figure 6.18 Initial Work Load Distribution-2

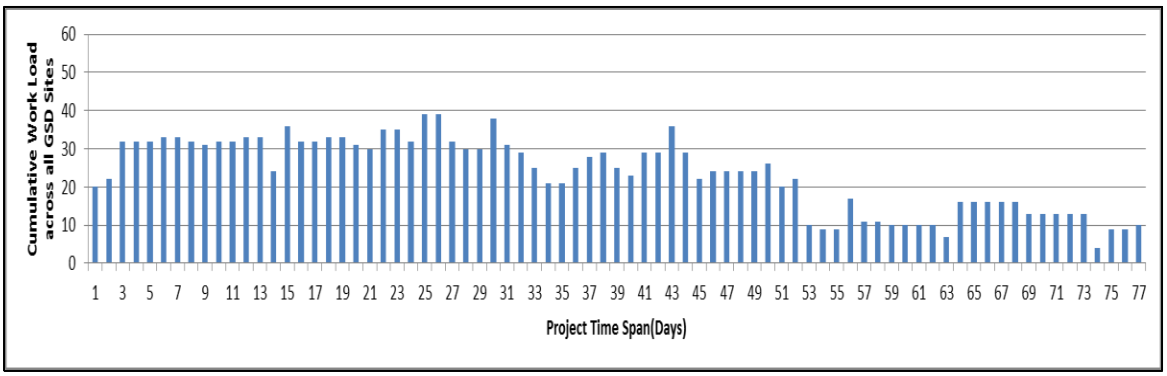

Figure 6.19 Optimal Work Load Distribution (Resource leveling)-2

Figure 6.19 shows the optimal work load matching as well as resource leveling i.e. try to distribute the work load equally across whole project time span. We can observe that after optimization the work load is equally distributed to some extent. The cumulative work load across all GSD sites on a particular day is between 20 man hours for most of the days in project time span. Cumulative work load for first 52 days in between 20 and 30 days and for remaining days cumulative work load is almost 10 man hours. Another major improvement in work load distribution is that now we have no particular days without any workload.

CHAPTER 7

DISCUSSION AND LIMITATIONS

This chapter explains some potential limitations of this work and conclude the work with potential future research directions. Section 7.2 explain the threats to validity and section **Error! Reference source not found.** conclude this work and also propose some future esearch possibilities.

7.1. Discussion

In this study first we perform Systematic Literature Review (SLR) and identify 11 factors that influence task allocation process in GSD projects. After that we conduct questionnaire survey from industry experts for factors identified from SLR process. In questionnaire survey the respondent is also ask to rank the positively impact factor based on relative importance with other positively impact factors.

Table 7.1 Relative importance of Factors based on Industrial Data

Factors	**Occurrence in survey (n=62)**							
	Relative Importance					**Relative importance for above 90% positive response**		
	VI	I	MI	SWI	NVI	Importance (%)	Import. Rank	+ve rank
Site Technical Expertise	31	24	2	2	2	90	3	1
Requirements Stability	36	16	6	1	1	87	5	2
Product Architecture	40	15	2	1	0	95	1	3
Task Dependency	35	19	5	0	0	91	2	3
Resource Cost	27	25	5	1	0	89	4	4

Table 7.1 shows the relative importance results based on industrial data. Five ranking scale is used for each factor, 'very important', important', 'moderately important', 'somewhat

important' and not important. Here we select only those factors that have more than 90% positive response from experts and related rank is shown in column 'positive rank'. The importance column shows the relative importance for that factor and relative rank column shows the rank based on relative importance percentage.

This analysis benefits to industry experts for GSD projects, we can observe that site technical expertise have positive response rank one and relative importance rank three, that shows the importance of this factor for task allocation process in GSD projects. Same can be apply for factor task dependency whose positive rank is three and relative importance rank is two. These two factors can easily be applied during task allocation process. We also use these two factor in proposed model and results shows considerable improvements and insure project completion on deadline also data is easily available for these factors. The other factor are also very important as result shows but all are subjective, we don't have such metric that can be used to measure these factors.

7.2. Threats to validity

In this section we discuss four major areas of threats to validity, namely Construct Validity, Internal Validity, External Validity and conclusion validity.

(i). Construct Validity

Construct validity deals with the concepts that are used in this study. Task allocation, factors and Global software development are the main construct that used in this study. For all these three concepts we use the actual terms and their synonyms i.e. first concept 'task allocation' and their synonyms 'work packages' and 'work distribution' etc., second concept ‘factors’ and their synonyms ‘causes’, ‘agents’ and ‘elements’ etc. and third concept ‘global software development’ and their synonyms ‘distributed software development’ and ‘global software engineering’ etc.in order to make sure that selected studies are related to these concepts. We could not perform a comprehensive manual search related to SLR due to the fact there are no platform (conference/journal) specifically focused on the joint use of these concepts. In order to mitigate this threat, we include the general intervention term ‘‘work distribution” along with ‘‘distributed software engineering’’ in the terms for the search in five reputable database.

(ii). Internal Validity

Paper selection and data extraction process may be influenced by some subjective decisions, as most of the papers does not clearly describe the study goals and results and reader have to make some assumptions from results description. In order to minimize this influence, paper selection and data extraction are carried out in iterative manner and by a group of researchers collaboratively. This collaboration of researchers and reviewers helps us to minimize the personal biasness effect and mitigate this threat.

(iii). External Validity

The inherent limitation of empirical studies lies in their external validity because of difficulty is achieving a true random sample of participants [103]. In our study, we mitigate possible bias related to sampling by using a variant of snowball sampling technique where key participants serve as contact points in the organizations involved[104]. We also provide definitions of key terms to avoid any misunderstanding from participants. Furthermore, we insure that all the potential participants have relevant experience in the development of global software systems. We believe that the results of the study are at least a reasonable indicator of practices in Asian companies developing global solutions.

(iv). Conclusion Validity

Papers selection, inclusion and exclusion criteria can depict the threat of conclusion validity because based on this process some papers may be excluded that should be included. In order to mitigate this threat we conducted the selection process individually and then selected the final studies based input from my supervisor (second reviewer).

CHAPTER 8

CONCLUSION AND FUTURE WORK

GSD has become a popular software development approach due to different benefits such as reduced development costs, access to a larger talent pool and follow-the-sun development. As GSD is expanding, we were motivated to identify task allocation criteria used in GSD projects.

In this study we performed an SLR to explore the factors that influence task allocation in GSD. Our main objective is to improve the task allocation process in GSD projects. Finally Thirty-eight primary studies were included. Through this SLR, we identified eleven factors of task allocation in GSD. The top ranked factors are: 'site technical expertise', 'time zone difference', 'resource cost', 'task dependency', 'vendor reliability', 'task size' and 'vendor maturity level'. After that we validate literature findings from industry experts. We applied statistical techniques like Chi-Square test, Fisher Exact test to analyze literature and industrial data based on client vendor perspective. We also compare literature findings with industrial expert's opinion using g statistical techniques. However literature lacks in terms of proposed solutions and tool usage for this problem.

Additionally, based on the literature findings, industrial data and best practices we implement task allocation model for GSD environment using genetic algorithm. We run our model on real time case study Life Cycle Assessment (LCA) tool as well one random data instance from Project Scheduling Library (PSPLiB). We only consider two factors "Task Dependency" and Site Technical Expertise" for these problems and Results shows substantial improvement in GSD in term of product quality and project deadline requirement.

For future work there are some potential areas that needs investigation like, explore the dependency between identified task allocation factors, further industrial studies are required to validate our findings and to provide a set of best practices, which can be used to address these factors and includes factors other than site technical expertise and task dependency like resource cost for task allocation to geographically distributed sites.

APPENDIX

A. SLR Protocol

❖ Research Question

Formulating the research questions is an important part of systematic literature review. To identify the factors that influence the task allocation in global software development, we designed these two research questions.

RQ 1: What are the factors that influence task allocation decisions in global software development (GSD) projects?

RQ 2: What are the possible solutions for the identified task allocation factors in GSD organizations?

After constructing the research questions we need to construct the search terms that answer these questions discussed in next section.

2.1. Constructing search terms

To identify relevant papers across various literary databases, we divided our research questions into sub-parts and then searched the databases with each sub-part and their synonymic grammar. The following technique was used to divide the research question into sub-part. Each sub-part is grammatically named like population, intervention and outcomes to form the basis for the construction of suitable search terms. However, not all research questions require an intervention.

Population: Global software development projects, task allocation, task allocation and GSD.
Intervention: factors that influence, solutions
Outcomes of Relevance: Factors that influence the task allocation decisions in global software development, possible solutions for these factors.
Experimental Design: Empirical investigation, Empirical studies, expert's opinions and case studies.

An example of a research question including the above mentioned details is:

RQ1:
[What are the factors that influence?] ***INTERVENTION***
[Task allocation decisions] ***POPULATION***
in
[Global Software Development] ***OUTCOMES OF RELEVANCE***

Syntactical changes were made depending upon the database being used. The literature search will not be limited with respect to time period and will cover all the relevant papers which have been published until now.

❖ Search Strategy

- **Identifying the search terms for search Database**

This section outlines the strategy that will be used to search for primary studies. The strategy used to construct the search term is explained below:

a. By identifying population, intervention and outcome we can derive the major terms from the research questions.
b. For the derived major terms we can find the alternate spelling and synonyms.
c. We can then verify the above steps with matching the keywords from any relevant research paper.
d. Depending on the search Database we can use Boolean operator 'AND' for concatenation of the major term and Boolean operator 'OR' for the concatenation of alternative spellings and synonyms.

Results for a)

For each research question, we identify the major terms separately.
RQ1: Global Software Development, factors, influence, task allocation decisions.
RQ2: Solutions for factors in Global Software Development, task allocation.

Results for b)

Global Software Development:
"Global software development projects" OR "global project management" OR "GSD" OR "Global Software Development" OR "Offshore software development" OR "offshore Outsourcing" OR "distributed software development" OR "Global Software Engineering " OR " Distributed Software Engineering " OR "GSE"

Factors:
"Factors" OR "causes" OR "agents" OR "elements" OR "aspects" OR "determinants" OR "constituents" OR "ingredients".

Task Allocation:
"Task allocation" OR "work distribution" OR "work assignment" OR "task assignment" OR "work allocation"
Solutions:
"Solutions" OR "answers" OR "Keys" OR "Strategy" OR "Processes" OR "Practices".

Results for c)

Global Software Development, factors, Solutions, task allocation.

Results for d)

The search strings for the specific electronic Databases are given below with the screenshots present in Appendix A respectively for each database.

RQ1)

((("Factors" OR "causes" OR "agents" OR "elements" OR "aspects" OR "determinants" OR

"constituents" OR "ingredients") AND "Task allocation" OR "work distribution" OR "work assignment" OR "task assignment" OR "work allocation") AND "Global software development projects" OR "global project management" OR "GSD" OR "Global Software Development" OR "Offshore software development" OR "Offshore Outsourcing" OR "distributed software development" OR "Global Software Engineering " OR " Distributed Software Engineering " OR "GSE"
)

IEEExplore- 881 Proper Results returned

RQ2)
(((("Solutions" OR "answers" OR "Keys" OR "Strategy" OR "Processes" OR "Practices") AND "Factors" OR "causes" OR "agents" OR "elements" OR "aspects" OR "determinants" OR "constituents" OR "ingredients") AND "Task allocation" OR "work distribution" OR "work assignment" OR "task assignment" OR "work allocation") AND "Global software development projects" OR "global project management" OR "GSD" OR "Global Software Development" OR "Offshore software development" OR "Offshore Outsourcing" OR "distributed software development" OR "Global Software Engineering " OR " Distributed Software Engineering " OR "GSE")

IEEExplore- 881 Proper Results returned

RQ1)
"Factors" OR "causes" OR "agents" OR "elements" OR "aspects" OR "determinants" OR "constituents" OR "ingredients" in All Fields AND "Task allocation" OR "work distribution" OR "work assignment" OR "task assignment" OR "work allocation" in All Fields And "Global software development projects" OR "global project management" OR "GSD" OR "Global Software Development" OR "Offshore software development" OR "distributed software development" OR "offshore outsourcing" OR "Global Software Engineering " OR " Distributed Software Engineering " OR "GSE" in All Fields

John Wiley Online Library- 35 Proper Results returned

RQ2)
"Solutions" OR "answers" OR "Keys" OR "Strategy" OR "Processes" OR "Practices" in All Fields "Factors" OR "causes" OR "agents" OR "elements" OR "aspects" OR "determinants" OR "constituents" OR "ingredients" in All Fields AND "Task allocation" OR "work distribution" OR "work assignment" OR "task assignment" OR "work allocation" in All Fields And "Global software development projects" OR "global project management" OR "GSD" OR "Global Software Development" OR "Offshore software development" OR "distributed software development" OR "offshore outsourcing" OR "Global Software Engineering " OR " Distributed Software Engineering " OR "GSE" in All Fields

John Wiley Online Library- 34 Proper Results returned

RQ1)
("Factors" OR "causes" OR "agents" OR "elements" OR "aspects" OR "determinants" OR "constituents" OR "ingredients") and ("task allocation" OR "work distribution" OR "work assignments" OR "task assignment" OR "work allocation") and ("global software development" OR "global project management" OR "GSD" OR "Global Software Development" OR "Offshore software development" OR "distributed software development" OR "offshore outsourcing" OR

"Global Software Engineering " OR " Distributed Software Engineering " OR "GSE")[All Sources(Computer Science)]
Science Direct- 32 proper results returned

RQ2)
("Solutions" OR "answers" OR "Keys" OR "Strategy" OR "Processes" OR "Practices") and ("Factors" OR "causes" OR "agents" OR "elements" OR "aspects" OR "determinants" OR "constituents" OR "ingredients") and ("task allocation" OR "work distribution" OR "work assignments" OR "task assignment" OR "work allocation") and ("global software development" OR "global project management" OR "GSD" OR "Global Software Development" OR "Offshore software development" OR "distributed software development" OR "offshore outsourcing" OR "Global Software Engineering " OR " Distributed Software Engineering " OR "GSE")[All Sources(Computer Science)]
Science Direct- 32 proper results returned

RQ1)
'(Factors OR causes OR agents OR elements OR aspects OR determinants OR constituents OR ingredients) AND (task allocation OR work distribution OR work assignments OR task assignment OR work allocation) AND ("global software development" OR "global project management" OR "GSD" OR "Global Software Development" OR "Offshore software development" OR "distributed software development" OR "offshore outsourcing" OR "Global Software Engineering " OR " Distributed Software Engineering " OR "GSE")' within Computer Science
SpringerLink- 218 proper results returned

RQ2)
'("Solutions" OR "answers" OR "Keys" OR "Strategy" OR "Processes" OR "Practices") AND (Factors OR causes OR agents OR elements OR aspects OR determinants OR constituents OR ingredients) AND (task allocation OR work distribution OR work assignments OR task assignment OR work allocation) AND ("global software development" OR "global project management" OR "GSD" OR "Global Software Development" OR "Offshore software development" OR "distributed software development" OR "offshore outsourcing" OR "Global Software Engineering " OR " Distributed Software Engineering " OR "GSE")' within Computer Science
SpringerLink- 216 proper results returned

RQ1)

((Factors or causes or agents or elements or aspects or determinants or constituents or ingredients) and ("Task allocation" or "work distribution" or "work assignments" or "task assignment" or "work allocation") and ("global software development" or "Global software development projects" or "global project management" or "GSD" or "Global Software Development" or "Offshore software development" or "distributed software development" or "offshore outsourcing" or "Global Software Engineering " or " Distributed Software Engineering " or "GSE"))
ACM- 123 proper results returned

RQ2)
(("Solutions" or "answers" or "Keys" or "Strategy" or "Processes" or "Practices") and (Factors or causes or agents or elements or aspects or determinants or constituents or ingredients) and ("Task allocation" or "work distribution" or "work assignments" or "task assignment" or "work allocation")

and ("global software development" or "Global software development projects" or "global project management" or "GSD" or "Global Software Development" or "Offshore software development" or "distributed software development" or "offshore outsourcing" or "Global Software Engineering " or " Distributed Software Engineering " or "GSE"))

ACM- 101 proper results returned

B. Task Allocation in Global Software Development Survey

Dear Sir/Madam

We should like to invite you to participate in the research project "Task Allocation in Global Software Development (GSD); An Empirical Study" being conducted by Sajid Anwer, Department of Information and Computer Sciences King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia, for the purpose of his Master degree (MS). The primary objective of this research is to assist Global Software Organizations (GSD) organizations in better understanding task allocation and work distribution activity in globally distrusted development projects. One critical factor of GSD for successful project completion with in specified time and budget, is the allocation of tasks as project managers not only need to consider her/his workforce but also need to take into the account the characteristics of the sites, their relationships and task characteristics. . As results become available, we shall provide you with copies of relevant papers and documents. We believe that these results will assist you in the design of suitable criteria for of task allocation GSD projects.

All information gathered from the questionnaire is for research purposes only. Such information will be treated in the STRICTEST CONFIDENCE and any publication from this study will present information in aggregate form such that individual organisations or individual respondents participating in the research cannot be identified. You can withdraw your participation at any time during this project. In addition only the supervisory team and Mr Sajid will have access to the data.

You can contact Sajid Anwer at +966591488381 or email g201303950@kfupm.edu.sa or his supervisor Dr Sajjad Mahmood at +966-3-860-7698 or email smahmood@kfupm.edu.sa if you have any concerns about the research. You are free to withdraw your participation from this research project at any time you wish and without giving a reason. We would appreciate your participation in this research.

Yours sincerely,
Dr. Sajjad Mahmood
Assistant Professor
KFUPM, Dhahran, Saudi Arabia

*Required

Section one: Practitiner's Detail

Position/Job Title *

Experience in years (Your experience in software outsourcing) *

Address *

Email *

Company,s country in which it is located? *

What is primary business function of your company? (you may tick more than one option) *

- ☐ In-house development
- ☐ Outhouse development
- ☐ Other:

What is the scope of your company? (Please tick as appropriate) *

- ☐ National
- ☐ Multinational
- ☐ Don't Know
- ☐ Other:

What type of Project Management Model typically used in your organization? *

☐ Distributed Project Management with Local Coordinators-- All or most of the team members report to local coordinators, who are responsible for the planning and execution of sub-projects or work packages.

☐ Distributed Project Management with Functional Coordinators-- Members related to certain functionality i.e. testing reported to functional coordinators, who are responsible for required functionality and might be located in different geographical areas.

Approximately how many staff are employed by your company? (Please tick as appropriate) *

- ☐ Less than 20
- ☐ 20-199
- ☐ Greater than 200
- ☐ Not Sure

Approximately how many staff are employed directly in the production/maintenance of software? (Please tick as appropriate) *

- ☐ Less than 20
- ☐ 20-199
- ☐ Greater than 200
- ☐ Not sure

What type of systems are your company concerned with? (You may tick more than one) *

- ☐ Safety Critical
- ☐ Business Systems
- ☐ Telecommunications
- ☐ Real Time systems
- ☐ Data processing
- ☐ System Software
- ☐ Windows based
- ☐ Embedded Systems
- ☐ Other: ____

Section 2

2.1. Evaluation of factors that influence task allocation in Global Software Development (GSD) Projects Identified by the Systematic Literature Review. The objective of this question is to identify factors influencing task allocation in GSD projects. Please cross the appropriate box based on your experience in the development of GSD projects.

Factors influence task allocation in GSD Projects *

	Strongly Agree	Agree	Disagree	Strongly Disagree
Site Technical Expertise	○	○	○	○
Task Dependency	○	○	○	○
Time Zone Difference	○	○	○	○
Vendor Reliability	○	○	○	○
Resource Cost	○	○	○	○
Task Size	○	○	○	○
Vendor Maturity Level	○	○	○	○
Local Government Regulations	○	○	○	○
Intellectual Property Ownership	○	○	○	○
Infrastructure Difference	○	○	○	○
Product Architecture	○	○	○	○
Requirements Stability	○	○	○	○

Section 2.2

For 'Strongly Agree' or 'Agree' factors, please rank their respective importance on a scale as follows: 1. Very important 2. Important 3. Moderately important 4. Somewhat important 5. Not very important

*

	Very Important	Important	Moderately Important	Somewhat Important	Not Very Important
Site Technical Expertise					
Task Dependency					
Time Zone Difference					
Vendor Reliability					
Resource Cost					
Task Size					
Vendor Maturity Level					
Local Government Regulations					
Intellectual Property Ownership					
Infrastructure Difference					
Product Architecture					

	Very Important	Important	Moderately Important	Somewhat Important	Not Very Important
Requirements Stability	○	○	○	○	○

2.3. List best practices and techniques being used at your organization during the task allocation decisions in GSD projects.

C. Client Vendor Analysis: Literature Findings

Paper_ID	Dependency	Size	Time	Expertise	Goverment	Maturity	Reliability	Requirement	Architecture	IP	Cost	clientorvendor
1.00	1.00	1.00	2.00	1.00	2.00	2.00	2.00	1.00	2.00	1.00	2.00	2.00
2.00	1.00	2.00	1.00	1.00	1.00	1.00	1.00	2.00	2.00	2.00	1.00	1.00
3.00	1.00	1.00	2.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
4.00	2.00	2.00	1.00	1.00	1.00	1.00	1.00	2.00	2.00	2.00	2.00	1.00
5.00	1.00	2.00	2.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
6.00	1.00	2.00	1.00	1.00	2.00	1.00	1.00	2.00	2.00	2.00	1.00	1.00
7.00	1.00	2.00	1.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
8.00	1.00	2.00	1.00	1.00	2.00	1.00	1.00	1.00	2.00	2.00	1.00	1.00
9.00	1.00	2.00	1.00	1.00	1.00	1.00	1.00	2.00	2.00	2.00	1.00	1.00
10.0	1.00	2.00	1.00	2.00	2.00	2.00	1.00	2.00	2.00	2.00	1.00	1.00
11.0	1.00	1.00	1.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	1.00	1.00
12.0	1.00	2.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	1.00	1.00
13.0	1.00	1.00	1.00	2.00	2.00	1.00	1.00	1.00	2.00	2.00	1.00	1.00
14.0	1.00	2.00	1.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	1.00	1.00
15.0	1.00	1.00	1.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
16.0	1.00	2.00	1.00	2.00	2.00	2.00	1.00	2.00	1.00	2.00	2.00	1.00
17.0	1.00	1.00	1.00	1.00	2.00	2.00	1.00	2.00	2.00	2.00	1.00	1.00

18.0	2.00	2.00	2.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	1.00
19.0	2.00	2.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	1.00
20.0	1.00	1.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
21.0	2.00	1.00	1.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
22.0	2.00	2.00	2.00	1.00	2.00	1.00	1.00	2.00	2.00	2.00	2.00	1.00
23.0	2.00	2.00	2.00	2.00	2.00	2.00	1.00	2.00	2.00	2.00	2.00	1.00
24.0	2.00	2.00	1.00	2.00	2.00	2.00	1.00	2.00	1.00	2.00	1.00	1.00
25.0	2.00	1.00	2.00	2.00	2.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00
26.0	2.00	1.00	1.00	1.00	1.00	2.00	2.00	2.00	2.00	2.00	1.00	2.00
27.0	2.00	2.00	1.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	1.00	1.00
28.0	2.00	1.00	1.00	1.00	1.00	2.00	2.00	2.00	2.00	2.00	1.00	2.00
29.0	2.00	2.00	1.00	2.00	2.00	2.00	1.00	2.00	1.00	2.00	1.00	2.00
30.0	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	1.00
31.0	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	1.00	2.00
32.0	2.00	2.00	2.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	1.00	1.00
33.0	2.00	2.00	2.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	1.00	1.00
34.0	2.00	2.00	2.00	1.00	2.00	2.00	1.00	2.00	2.00	2.00	2.00	2.00
35.0	2.00	2.00	1.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
36.0	2.00	2.00	1.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
37.0	2.00	2.00	2.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	1.00
38.0	2.00	2.00	2.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	1.00

D. Client Vendor Analysis: Questionnaire Survey

Number	Dependency	Task Size	Time Zone	Cost	Requirements	Architecture	Government	IP	Expertise	Maturity	Reliability	Client or vendor
1.00	1.00	3.00	1.00	2.00	2.00	2.00	3.00	2.00	1.00	1.00	1.00	1.00
2.00	1.00	1.00	2.00	1.00	1.00	1.00	2.00	2.00	1.00	2.00	1.00	1.00
3.00	1.00	1.00	2.00	1.00	1.00	1.00	3.00	1.00	2.00	1.00	1.00	1.00
4.00	2.00	3.00	4.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	1.00	1.00
5.00	1.00	3.00	2.00	1.00	2.00	1.00	3.00	2.00	2.00	1.00	1.00	1.00
6.00	1.00	3.00	1.00	1.00	1.00	1.00	4.00	1.00	1.00	1.00	1.00	1.00
7.00	1.00	1.00	1.00	1.00	1.00	2.00	1.00	1.00	1.00	2.00	1.00	1.00
8.00	2.00	3.00	2.00	2.00	1.00	1.00	2.00	2.00	2.00	2.00	1.00	1.00

9.00	2.00	2.00	2.00	1.00	2.00	2.00	3.00	2.00	2.00	1.00	1.00	1.00
10.00	1.00	2.00	2.00	1.00	1.00	1.00	2.00	2.00	2.00	1.00	2.00	1.00
11.00	1.00	2.00	2.00	1.00	1.00	1.00	3.00	1.00	1.00	1.00	1.00	1.00
12.00	2.00	2.00	2.00	2.00	2.00	2.00	1.00	2.00	2.00	2.00	2.00	1.00
13.00	2.00	2.00	2.00	1.00	2.00	3.00	3.00	3.00	2.00	1.00	2.00	1.00
14.00	1.00	1.00	2.00	1.00	1.00	1.00	3.00	1.00	1.00	1.00	1.00	1.00
15.00	1.00	1.00	2.00	1.00	1.00	1.00	2.00	2.00	1.00	2.00	1.00	1.00
16.00	1.00	1.00	2.00	1.00	1.00	1.00	3.00	1.00	2.00	1.00	2.00	1.00
17.00	2.00	2.00	4.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	1.00
18.00	2.00	2.00	2.00	2.00	2.00	2.00	1.00	2.00	2.00	2.00	2.00	1.00
19.00	1.00	1.00	1.00	1.00	1.00	2.00	1.00	1.00	1.00	2.00	2.00	1.00
20.00	2.00	2.00	2.00	2.00	1.00	1.00	2.00	2.00	2.00	2.00	2.00	1.00
21.00	2.00	2.00	3.00	3.00	2.00	2.00	2.00	2.00	2.00	2.00	3.00	2.00
22.00	1.00	1.00	2.00	3.00	2.00	2.00	2.00	1.00	2.00	2.00	2.00	2.00
23.00	1.00	1.00	2.00	1.00	1.00	1.00	2.00	2.00	1.00	1.00	4.00	2.00
24.00	1.00	2.00	1.00	1.00	2.00	2.00	3.00	1.00	1.00	2.00	3.00	2.00
25.00	1.00	2.00	3.00	1.00	1.00	1.00	1.00	2.00	2.00	2.00	2.00	2.00
26.00	2.00	2.00	1.00	2.00	1.00	1.00	2.00	2.00	2.00	2.00	3.00	2.00
27.00	1.00	2.00	1.00	1.00	2.00	1.00	4.00	4.00	2.00	3.00	2.00	2.00
28.00	2.00	2.00	2.00	2.00	1.00	2.00	3.00	3.00	2.00	2.00	3.00	2.00
29.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	1.00	2.00
30.00	1.00	2.00	3.00	2.00	1.00	2.00	1.00	2.00	1.00	3.00	3.00	2.00
31.00	1.00	1.00	1.00	2.00	3.00	3.00	3.00	3.00	3.00	3.00	1.00	2.00
32.00	2.00	1.00	2.00	2.00	2.00	2.00	1.00	2.00	2.00	3.00	2.00	2.00
33.00	2.00	2.00	2.00	1.00	1.00	1.00	2.00	1.00	1.00	2.00	1.00	2.00
34.00	2.00	2.00	3.00	2.00	2.00	2.00	3.00	3.00	1.00	2.00	2.00	2.00
35.00	1.00	2.00	2.00	2.00	4.00	3.00	3.00	1.00	1.00	1.00	2.00	2.00
36.00	3.00	1.00	2.00	2.00	2.00	2.00	2.00	3.00	2.00	2.00	1.00	2.00
37.00	2.00	1.00	2.00	2.00	2.00	1.00	2.00	2.00	1.00	2.00	1.00	2.00
38.00	2.00	1.00	3.00	2.00	2.00	2.00	3.00	2.00	2.00	2.00	3.00	2.00
39.00	2.00	1.00	2.00	3.00	2.00	2.00	3.00	2.00	1.00	2.00	3.00	2.00
40.00	1.00	1.00	3.00	2.00	1.00	1.00	2.00	2.00	2.00	1.00	1.00	2.00
41.00	2.00	2.00	3.00	2.00	1.00	2.00	3.00	2.00	2.00	1.00	1.00	2.00
42.00	1.00	1.00	1.00	1.00	2.00	1.00	3.00	2.00	2.00	2.00	3.00	2.00
43.00	1.00	1.00	2.00	1.00	1.00	2.00	2.00	1.00	2.00	2.00	1.00	2.00
44.00	1.00	1.00	2.00	1.00	1.00	1.00	1.00	2.00	1.00	1.00	1.00	2.00
45.00	2.00	1.00	3.00	2.00	2.00	2.00	2.00	2.00	2.00	1.00	1.00	2.00
46.00	1.00	1.00	1.00	1.00	1.00	2.00	2.00	1.00	2.00	1.00	1.00	2.00
47.00	1.00	1.00	2.00	1.00	2.00	1.00	2.00	2.00	1.00	1.00	1.00	2.00

48.00	1.00	1.00	4.00	2.00	1.00	1.00	3.00	2.00	2.00	3.00	3.00	2.00
49.00	3.00	1.00	1.00	2.00	1.00	1.00	2.00	2.00	1.00	3.00	1.00	2.00
50.00	2.00	1.00	2.00	2.00	2.00	2.00	1.00	2.00	2.00	2.00	2.00	2.00
51.00	2.00	3.00	2.00	1.00	1.00	1.00	2.00	1.00	1.00	1.00	1.00	2.00
52.00	2.00	1.00	3.00	2.00	2.00	2.00	3.00	3.00	1.00	2.00	2.00	2.00
53.00	2.00	2.00	3.00	2.00	2.00	2.00	2.00	2.00	2.00	1.00	1.00	2.00
54.00	1.00	2.00	1.00	1.00	1.00	2.00	2.00	1.00	2.00	1.00	1.00	2.00
55.00	1.00	1.00	4.00	2.00	1.00	1.00	3.00	2.00	2.00	3.00	3.00	2.00
56.00	3.00	2.00	1.00	2.00	1.00	1.00	2.00	2.00	1.00	3.00	1.00	2.00
57.00	2.00	2.00	2.00	3.00	2.00	2.00	3.00	2.00	1.00	2.00	3.00	2.00
58.00	2.00	2.00	3.00	2.00	2.00	2.00	3.00	3.00	1.00	2.00	2.00	2.00
59.00	1.00	1.00	3.00	2.00	1.00	1.00	2.00	2.00	2.00	1.00	1.00	2.00
60.00	2.00	2.00	3.00	2.00	1.00	2.00	3.00	2.00	2.00	1.00	1.00	2.00
61.00	1.00	1.00	1.00	1.00	2.00	1.00	3.00	2.00	2.00	2.00	3.00	2.00
62.00	1.00	3.00	2.00	1.00	1.00	2.00	2.00	1.00	2.00	2.00	1.00	2.00

References

[1] T. Kern and L. Willcocks, "Exploring information technology outsourcing relationships: theory and practice," *The Journal of Strategic Information Systems,* vol. 9, pp. 321-350, 2000.

[2] U. UNCTAD, "World investment report: The shift towards services," *New York: UN,* 2004.

[3] D. Bradstreet, "Dun & Bradstreet's Barometer of Global Outsourcing," *Dun & Bradstreet,* 2000.

[4] S. Islam, M. M. A. Joarder, and S. H. Houmb, "Goal and Risk Factors in Offshore Outsourced Software Development From Vendor's Viewpoint," in *IEEE International Conference on Global Software Engineering, ICGSE09*, Limerick, Ireland, 2009, pp. 347-352.

[5] S. U. Khan, M. Niazi, and R. Ahmad, "Critical success factors for offshore software development outsourcing vendors: A systematic literature review," in *Global Software Engineering, 2009. ICGSE 2009. Fourth IEEE International Conference on*, 2009, pp. 207-216.

[6] A. A. Bush, A. Tiwana, and H. Tsuji, "An empirical investigation of the drivers of software outsourcing decisions in Japanese organizations," *Information and Software Technology,* vol. 50, pp. 499-510, 2008.

[7] A. von Stetten, D. Beimborn, E. Kuznetsova, and B. Moos, "The impact of cultural differences on IT nearshoring risks from a German perspective," in *System Sciences (HICSS), 2010 43rd Hawaii International Conference on*, 2010, pp. 1-10.

[8] B. Shao and J. S. David, "The impact of offshore outsourcing on IT workers in developed countries," *Communications of the ACM,* vol. 50, pp. 89-94, 2007.

[9] H. Holmstrom, E. Ó. Conchúir, P. J. Agerfalk, and B. Fitzgerald, "Global software development challenges: A case study on temporal, geographical and socio-cultural distance," in *Global Software Engineering, 2006. ICGSE'06. International Conference on*, 2006, pp. 3-11.

[10] S. U. Khan, M. Niazi, and R. Ahmad, "Barriers in the selection of offshore software development outsourcing vendors: An exploratory study using a systematic literature review," *Information and Software Technology,* vol. 53, pp. 693-706, 2011.

[11] M. Belcourt, "Outsourcing—The benefits and the risks," *Human resource management review,* vol. 16, pp. 269-279, 2006.

[12] S. U. Khan, M. Niazi, and R. Ahmad, "Factors influencing clients in the selection of offshore software outsourcing vendors: An exploratory study using a systematic literature review," *Journal of systems and software,* vol. 84, pp. 686-699, 2011.

[13] K. Nam, S. Rajagopalan, H. R. Rao, and A. Chaudhury, "A two-level investigation of information systems outsourcing," *Communications of the ACM,* vol. 39, pp. 36-44, 1996.
[14] B. A. Aubert, S. Dussault, M. Patry, and S. Rivard, "Managing the risk of IT outsourcing," in *Systems Sciences, 1999. HICSS-32. Proceedings of the 32nd Annual Hawaii International Conference on*, 1999, p. 10 pp.
[15] J. A. Blackley and J. Leach, "Security considerations in outsourcing IT services," *Information Security Technical Report,* vol. 1, pp. 11-17, 1996.
[16] J. Pyysiäinen, "Building trust in global inter-organizational software development projects: problems and practices," in *International Workshop on Global Software Development*, 2003, pp. 69-74.
[17] R. Gonzalez, J. Gasco, and J. Llopis, "Information systems outsourcing risks: a study of large firms," *Industrial management & Data systems,* vol. 105, pp. 45-62, 2005.
[18] W. Kobitzsch, D. Rombach, and R. L. Feldmann, "Outsourcing in India [software development]," *Software, IEEE,* vol. 18, pp. 78-86, 2001.
[19] J. Barthelemy, "The hidden costs of IT outsourcing," *Sloan Management Review,* vol. 42, pp. 60-69, 2001.
[20] B. Kitchenham and C. Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering," Keele University and Durham University Joint Report EBSE 2007-001, 2007.
[21] S. Keele, "Guidelines for performing systematic literature reviews in software engineering," Technical report, EBSE Technical Report EBSE-2007-012007.
[22] P. Lago, H. Muccini, and M. A. Babar, "Developing a course on designing software in globally distributed teams," in *Global Software Engineering, 2008. ICGSE 2008. IEEE International Conference on*, 2008, pp. 249-253.
[23] T. Kern and L. Willcocks, "Exploring information technology outsourcing relationships: theory and practice," *Journal of Strategic Information Systems,* vol. 9, pp. 321-350, 2000.
[24] S. U. Khan, "Software outsourcing vendors' readiness model (SOVRM)," Keele University, 2011.
[25] P. A. Laplante, T. Costello, P. Singh, S. Bindiganavile, and M. Landon, "The who, what, why, where, and when of IT outsourcing," *IT professional,* vol. 6, pp. 19-23, 2004.
[26] B. Shao, David, J.S., "The impact of offshore outsourcing on IT workers in developed countries.," *Communications of the ACM,* vol. 50, pp. 89 - 94, 2007.
[27] A. Stetten, v., D. Beimborn, E. Kuznetsova, and B. Moos, "The Impact of Cultural Differences on IT Nearshoring Risks from a German Perspective," in *in Proceedings of the 43rd IEEE Hawaii International Conference on System Sciences*, 2010.

[28] M. Hopkins, "The antihero's guide to the new economy," *INC-BOSTON MA-,* vol. 20, pp. 36-49, 1998.
[29] E. Carmel and P. Abbott, "Configurations of global software development: offshore versus nearshore," in *Proceedings of the 2006 international workshop on Global software development for the practitioner*, 2006, pp. 3-7.
[30] D. Smite, C. Wohlin, T. Gorscheck, and R. Feldt, "Empirical evidence in global software engineering: a systematic review," *Empirical Software Engineering,* vol. 15, pp. 91-118, 2010.
[31] N. V. Oza, T. Hall, A. Rainer, and S. G. Grey, "Trust in software outsourcing relationships: An empirical investigation of Indian software companies," *Information & Software Technology,* vol. 48, pp. 345-354, 2006.
[32] M. D. Minevich and F.-J. Richter, *The global outsourcing report*: Palgrave Macmillan, 2006.
[33] M. J. Gallivan and W. Oh, "Analyzing IT outsourcing relationships as alliances among multiple clients and vendors," in *Systems Sciences, 1999. HICSS-32. Proceedings of the 32nd Annual Hawaii International Conference on*, 1999, p. 15 pp.
[34] L. Obal, "Microsourcing--using information technology to create unexpected work relationships and entrepreneurial opportunities: work in progress," in *Proceedings of the 2006 ACM SIGMIS CPR conference on computer personnel research: Forty four years of computer personnel research: achievements, challenges & the future*, 2006, pp. 60-62.
[35] H. Tsuji, A. Sakurai, K. Yoshida, A. Tiwana, and A. Bush, "Questionnaire-Based Risk Assessment Scheme for Japanese Offshore Software Outsourcing," in *SEAFOOD07, Springer*, 2007, pp. 114-127.
[36] L. McLaughlin, "An eye on India: Outsourcing debate continues.," *IEEE Software,* vol. 20, pp. 114-117, (May / June 2003) 2003.
[37] D. Damian, L. Izquierdo, J. Singer, and I. Kwan, "Awareness in the Wild: Why Communication Breakdowns Occur," in *International Conference on Global Software Engineering*, 2007, pp. 81-90.
[38] P. F. Takac, "Outsourcing: A key to controlling escalating IT costs," *International Journal of Technology Management,* vol. 9, pp. 139-155, 1994.
[39] D. Foote, "Recipe for offshore outsourcing failure: Ignore organization, people issues," *ABA Banking Journal,* vol. 96, pp. 56-59, 2004.
[40] W. R. King, "Outsourcing Becomes More Complex," *IS Management,* vol. 22, pp. 89-90, 2005.
[41] G. K. Hanssen, D. Smite, and N. B. Moe, "Signs of agile trends in global software engineering research: A tertiary study," in *Global Software Engineering Workshop (ICGSEW), 2011 Sixth IEEE International Conference on*, 2011, pp. 17-23.

[42] I. Richardson, V. Casey, F. McCaffery, J. Burton, and S. Beecham, "A process framework for global software engineering teams," *Information and Software Technology,* vol. 54, pp. 1175-1191, 2012.

[43] A. B. Marques, R. Rodrigues, and T. Conte, "Systematic literature reviews in distributed software development: A tertiary study," in *Global Software Engineering (ICGSE), 2012 IEEE Seventh International Conference on*, 2012, pp. 134-143.

[44] J. Kroll, S. I. Hashmi, I. Richardson, and J. L. Audy, "A Systematic Literature Review of Best Practices and Challenges in Follow-the-Sun Software Development," in *Global Software Engineering Workshops (ICGSEW), 2013 IEEE 8th International Conference on*, 2013, pp. 18-23.

[45] E. Carmel, J. A. Espinosa, and Y. Dubinsky, "" Follow the Sun" Workflow in Global Software Development," *Journal of Management Information Systems,* vol. 27, pp. 17-38, 2010.

[46] P. Mohapatra, P. Bjorndal, and K. Smiley, "Causal analysis of factors governing collaboration in global software development teams," in *Global Software Engineering (ICGSE), 2010 5th IEEE International Conference on*, 2010, pp. 128-132.

[47] A. Lamersdorf, J. Munch, and D. Rombach, "A survey on the state of the practice in distributed software development: Criteria for task allocation," in *Global Software Engineering, 2009. ICGSE 2009. Fourth IEEE International Conference on*, 2009, pp. 41-50.

[48] A. Lamersdorf, J. Munch, A. Fernández-del Viso Torre, and C. Rebate Sanchez, "A risk-driven model for work allocation in global software development projects," in *Global Software Engineering (ICGSE), 2011 6th IEEE International Conference on*, 2011, pp. 15-24.

[49] N. C. Narendra, K. Ponnalagu, N. Zhou, and W. M. Gifford, "Towards a Formal Model for Optimal Task-Site Allocation and Effort Estimation in Global Software Development," in *SRII Global Conference (SRII), 2012 Annual*, 2012, pp. 470-477.

[50] D. Wickramaarachchi and R. Lai, "A method for work distribution in Global Software Development," in *Advance Computing Conference (IACC), 2013 IEEE 3rd International*, 2013, pp. 1443-1448.

[51] M. Staples and M. Niazi, "Experiences using systematic review guidelines," *Journal of systems and software,* vol. 80, pp. 1425-1437, 2007.

[52] S. Mahdavi-Hezavehi, M. Galster, and P. Avgeriou, "Variability in quality attributes of service-based software systems: A systematic literature review," *Information and Software Technology,* vol. 55, pp. 320-343, 2013.

[53] T. Dybå and T. Dingsøyr, "Empirical studies of agile software development: A systematic review," *Information and Software Technology,* vol. 50, pp. 833-859, 2008.

[54] P. Achimugu, A. Selamat, R. Ibrahim, and M. N. r. Mahrin, "A systematic literature review of software requirements prioritization research," *Information and Software Technology,* vol. 56, pp. 568-585, 2014.
[55] D. Dermeval, J. Vilela, I. I. Bittencourt, J. Castro, S. Isotani, P. Brito, and A. Silva, "Applications of ontologies in requirements engineering: a systematic review of the literature," *Requirements Engineering,* pp. 1-33, 2015.
[56] W. Ding, P. Liang, A. Tang, and H. Van Vliet, "Knowledge-based approaches in software documentation: A systematic literature review," *Information and Software Technology,* vol. 56, pp. 545-567, 2014.
[57] C. B. Seaman, "Qualitative methods in empirical studies of software engineering," *Software Engineering, IEEE Transactions on,* vol. 25, pp. 557-572, 1999.
[58] S. Michael and B. Lewis, "Research practice: International handbooks of quantitative application in social sciences," ed: Sage, 1994.
[59] P. J. Stone, D. C. Dunphy, and M. S. Smith, "The General Inquirer: A Computer Approach to Content Analysis," 1966.
[60] J. W. Creswell, *Qualitative inquiry and research design: Choosing among five approaches*: Sage publications, 2012.
[61] N. Baddoo, "Motivators and de-motivators in software process improvement: an empirical study," University of Hertfordshire, 2001.
[62] P. Burnard, "A method of analysing interview transcripts in qualitative research," *Nurse education today,* vol. 11, pp. 461-466, 1991.
[63] O. F. Althuwaynee, B. Pradhan, H.-J. Park, and J. H. Lee, "A novel ensemble decision tree-based CHi-squared Automatic Interaction Detection (CHAID) and multivariate logistic regression models in landslide susceptibility mapping," *Landslides,* pp. 1-16, 2014.
[64] A. Lamersdorf and J. Münch, "A multi-criteria distribution model for global software development projects," *Journal of the Brazilian Computer Society,* vol. 16, pp. 97-115, 2010.
[65] A. Lamersdorf and J. Münch, "Model-based task allocation in distributed software development," in *Software Engineering Approaches for Offshore and Outsourced Development*, ed: Springer, 2010, pp. 37-53.
[66] A. Lamersdorf, J. Münch, and D. Rombach, "A decision model for supporting task allocation processes in global software development," in *Product-Focused Software Process Improvement*, ed: Springer, 2009, pp. 332-346.
[67] J. Münch and A. Lamersdorf, "Systematic task allocation evaluation in distributed software development," in *On the Move to Meaningful Internet Systems: OTM 2009 Workshops*, 2009, pp. 228-237.
[68] T. A. B. Pereira, V. S. dos Santos, B. L. Ribeiro, and G. Elias, "A recommendation framework for allocating global software teams in

software product line projects," in *Proceedings of the 2nd International Workshop on Recommendation Systems for Software Engineering*, 2010, pp. 36-40.
[69] Y. Ye, K. Nakakoji, and Y. Yamamoto, "Measuring and monitoring task couplings of developers and development sites in global software development," in *Software Engineering Approaches for Offshore and Outsourced Development*, ed: Springer, 2009, pp. 181-195.
[70] H. Hu, B. Xu, Y. Ling, X. Yang, Z. He, and A. Ma, "Micro-estimation Based Global Collaborative Task Arrangement in Distributed Software Design," in *Computer Supported Cooperative Work in Design IV*, ed: Springer, 2008, pp. 64-75.
[71] A. Gupta, I. Crk, and R. Bondade, "Leveraging temporal and spatial separations with the 24-hour knowledge factory paradigm," *Information Systems Frontiers,* vol. 13, pp. 397-405, 2011.
[72] S. Vathsavayi, O. Sievi-Korte, K. Koskimies, and K. Systä, "Planning Global Software Development Projects Using Genetic Algorithms," in *Search Based Software Engineering*, ed: Springer, 2013, pp. 269-274.
[73] A. Lamersdorf and J. Munch, "TAMRI: a tool for supporting task distribution in global software development projects," in *Global Software Engineering, 2009. ICGSE 2009. Fourth IEEE International Conference on*, 2009, pp. 322-327.
[74] J. Helming, H. Arndt, Z. Hodaie, M. Koegel, and N. Narayan, "Automatic assignment of work items," in *Evaluation of Novel Approaches to Software Engineering*, ed: Springer, 2011, pp. 236-250.
[75] M. Yilmaz and R. V. O'Connor, *A market based approach for resolving resource constrained task allocation problems in a software development process*: Springer, 2012.
[76] S. Doma, L. Gottschalk, T. Uehara, and J. Liu, "Resource allocation optimization for GSD projects," in *Computational Science and Its Applications–ICCSA 2009*, ed: Springer, 2009, pp. 13-28.
[77] N. Celik, S. Lee, E. Mazhari, Y.-J. Son, R. Lemaire, and K. G. Provan, "Simulation-based workforce assignment in a multi-organizational social network for alliance-based software development," *Simulation Modelling Practice and Theory,* vol. 19, pp. 2169-2188, 2011.
[78] P. Jalote and G. Jain, "Assigning tasks in a 24-h software development model," *Journal of Systems and Software,* vol. 79, pp. 904-911, 2006.
[79] A. Barcus and G. Montibeller, "Supporting the allocation of software development work in distributed teams with multi-criteria decision analysis," *Omega,* vol. 36, pp. 464-475, 2008.
[80] S. Abufardeh and K. Magel, "The impact of global software cultural and linguistic aspects on Global Software Development process (GSD): Issues

and challenges," in *New Trends in Information Science and Service Science (NISS), 2010 4th International Conference on*, 2010, pp. 133-138.
[81] A. Lamersdorf, J. Münch, A. F. Viso Torre, C. R. Sánchez, M. Heinz, and D. Rombach, "A rule-based model for customized risk identification and evaluation of task assignment alternatives in distributed software development projects," *Journal of Software: Evolution and Process,* vol. 24, pp. 661-675, 2012.
[82] S. o. Setamanit, W. Wakeland, and D. Raffo, "Using simulation to evaluate global software development task allocation strategies," *Software Process: Improvement and Practice,* vol. 12, pp. 491-503, 2007.
[83] M. Ruano-Mayoral, C. Casado-Lumbreras, H. Garbarino-Alberti, and S. Misra, "Methodological framework for the allocation of work packages in global software development," *Journal of Software: Evolution and Process,* vol. 26, pp. 476-487, 2014.
[84] D. K. Mak and P. Kruchten, "Task coordination in an agile distributed software development environment," in *Electrical and Computer Engineering, 2006. CCECE'06. Canadian Conference on*, 2006, pp. 606-611.
[85] A. Lamersdorf, J. Munch, and D. Rombach, "Towards a multi-criteria development distribution model: An analysis of existing task distribution approaches," in *Global Software Engineering, 2008. ICGSE 2008. IEEE International Conference on*, 2008, pp. 109-118.
[86] A. B. Marques, J. R. Carvalho, R. Rodrigues, T. Conte, R. Prikladnicki, and S. Marczak, "An Ontology for Task Allocation to Teams in Distributed Software Development," in *Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on*, 2013, pp. 21-30.
[87] M. Cataldo and S. Nambiar, "Quality in global software development projects: A closer look at the role of distribution," in *Global Software Engineering, 2009. ICGSE 2009. Fourth IEEE International Conference on*, 2009, pp. 163-172.
[88] J. Fernandez and M. Basavaraju, "Task allocation model in globally distributed software projects using genetic algorithms," in *Global Software Engineering (ICGSE), 2012 IEEE Seventh International Conference on*, 2012, pp. 181-181.
[89] D. K. Mak and P. B. Kruchten, "NextMove: A framework for distributed task coordination," in *Software Engineering Conference, 2007. ASWEC 2007. 18th Australian*, 2007, pp. 399-408.
[90] M. Ruano-Mayoral, R. Colomo-Palacios, J. M. Fernández-González, and Á. García-Crespo, "Towards a framework for work package allocation for GSD," in *On the Move to Meaningful Internet Systems: OTM 2011 Workshops*, 2011, pp. 200-207.

[91] S.-o. Setamanit, W. Wakeland, and D. Raffo, "Exploring the impact of task allocation strategies for global software development using simulation," in *Software Process Change*, ed: Springer, 2006, pp. 274-285.
[92] I. Richardson, V. Casey, J. Burton, and F. McCaffery, "Global software engineering: A software process approach," in *Collaborative Software Engineering*, ed: Springer, 2010, pp. 35-56.
[93] S. Imtiaz, "Architectural task allocation in distributed environment: a traceability perspective," in *Proceedings of the 34th International Conference on Software Engineering*, 2012, pp. 1515-1518.
[94] S. Deshpande, S. Beecham, and I. Richardson, "Global Software Development Coordination Strategies-A Vendor Perspective," in *New Studies in Global IT and Business Service Outsourcing*, ed: Springer, 2011, pp. 153-174.
[95] A. Lamersdorf, J. Munch, A. F.-d. V. Torre, C. R. Sánchez, and D. Rombach, "Estimating the effort overhead in global software development," in *Global Software Engineering (ICGSE), 2010 5th IEEE International Conference on*, 2010, pp. 267-276.
[96] R. D. Battin, R. Crocker, J. Kreidler, and K. Subramanian, "Leveraging resources in global software development," *Software, IEEE,* vol. 18, pp. 70-77, 2001.
[97] S. Goldman, J. Munch, and H. Holz, "A meta-model for distributed software development," in *Enabling Technologies: Infrastructure for Collaborative Enterprises, 1999.(WET ICE'99) Proceedings. IEEE 8th International Workshops on*, 1999, pp. 48-53.
[98] "http://academic.udayton.edu/gregelvers/psy216/spss/ttests.htm."
[99] "http://www.petronet.ir/documents/10180/2323250/advanced_project_management-_3ed_edition."
[100] "http://honestyets.pbworks.com/f/Project+Management+-+A+Systems+Approach+-+10thEd.pdf."
[101] "http://www.esu-services.ch/ourservices/case-studies/."
[102] "http://www.om-db.wi.tum.de/psplib/main.html."
[103] S. Mahmood, "Empirical study of software component integration process activities," *IET software,* vol. 7, pp. 65-75, 2013.
[104] B. Kitchenham and S. L. Pfleeger, "Principles of survey research: part 5: populations and samples," *ACM SIGSOFT Software Engineering Notes,* vol. 27, pp. 17-20, 2002.

Vitae

Name:	Sajid Anwer
Nationality:	Pakistan
Date of Birth:	1/2/1990
Email:	sajidanwer786@gmail.com
Address:	Building 810-113,KFUPM, Dhahran,31261,KSA

Academic Background:

Institute Name	**Degree**	**Duration**	**CGPA**
KFUPM, Dhahran, Saudi Arabia	MS(CS)	Sep, 2013 – May,2015	3.714
CIIT Lahore, Pakistan	BS(CS)	Sep, 2008 – June,2012	3.66 (Distinction)

Publications:

1. S. Mahmood, **S. Anwer,** M. Niazi, M. Alshayeb and I. Richardson "Task Allocation in Global Software Developemnt: Client Vendor Analysis" Information and Software Technologies Journal (Submitted)
2. **S. Anwer** and M. El-Attar, "An Evaluation of the Statechart Diagrams Visual Syntax," *International Conference on Information Science and Applications (ICISA), 2014*, pp. 1-4, 6-9 May 2014.
3. S. Mahmood, S. **Anwer,** W. Umer, M. Niazi, M. Alshayeb, "Towards Task Allocation in Global Software Development Projects." ICSEA, October 2014, France, ISBN: 978-1-61208-367-4.
4. S. Mahmood, S. Anwer, M. Niazi, M. Alshayeb, and I. Richardson, "Identifying the factors that influence task allocation in global software development: preliminary results," *presented at the Proceedings of the 19th International Conference on Evaluation and Assessment in Software Engineering*, Nanjing, China, April, 2015.