AALTO-YLIOPISTO
Perustieteiden korkeakoulu
Tietotekniikan koulutusohjelma

Sonja Tervola

Review and Analysis of Task Distribution Models as Mitigation Strategies for Challenges in Global Software Development

Kandidaatintyö

Espoo, 15.04.2022

Vastuuopettaja: Eero Hyvönen

Työnohjaaja: Fabian Fagerholm

Aalto-yliopisto Perustieteiden korkeakoulu Teknillistieteiden kandidaattiohjelma	KANDIDAATINTYÖN TIIVISTELMÄ

Tekijä: Sonja Tervola

Työn nimi: Review and Analysis of Task Distribution Models as Mitigation Strategies for Challenges in Global Software Development

Sivumäärä: 24	Päiväys: 15.04.2022	Julkaisukieli: englanti
Pääaine: Tietotekniikka		Pääaineen koodi: CS

Increasingly, software development is moving towards a more global approach: Global Software Development (GSD). This is because GSD can help to gain a competition advantage on the global market. In GSD, multinational development teams and people from different nations with different cultural backgrounds and geographical locations are involved in developing software. The thought major benefits of GSD include the division of human resources as well as communication technologies across different locations that can help lower costs and acquire higher efficiency through, for example, task optimization. Nonetheless, in reality GSD has shown to be technologically and organizationally more complicated as it is more difficult to manage, coordinate, and balance. GSD faces several challenges, mainly related to temporal, geographical and socio-cultural distances among the dispersed teams and people in distributed locations. However, efficient distribution of tasks can help to reduce the temporal, geographical and socio-cultural distances and solve the challenges of GSD. This thesis analyzes the challenges of GSD and suggests possible mitigation strategies. Task distribution models are identified and introduced as one of the most relevant and effective mitigation strategies for solving GSD challenges. Factors that influence task distribution activity in GSD as well as models that can help ameliorate task distribution management in GSD are further reviewed and analyzed.

Työnohjaaja(t): Fabian Fagerholm

Vastuuopettaja: Eero Hyvönen

Asiasanat: Review, Global Software Development, Challenges, Task Distribution, Mathematical models, Mitigation strategies

Table of Contents

1 Introduction

After centuries of globalization, the world is now much more connected than ever before. We have witnessed an ongoing trend towards the globalization of economics, businesses, technologies, and software. Economic forces are constantly pushing national markets to reach beyond their borders into global ones, creating new forms of competition and collaboration internationally. This effect of globalization has not only an impact on distribution and marketing, but, more importantly, on the fundamentals of business processes; design, construction, testing, and delivery to customers (Herbsleb, 2007).

Software industry has recognized the growing preponderance of globalization and the resulting competition, as it has, in recent years, attempted to understand the factors that contribute to successful multinational collaboration. For decades, companies have experimented with outsourcing, namely in the sphere of remotely located software development (Olson & Olson, 2000).

Herbsleb & Mockus (2003) define Global Software Development (GSD) as software development that is multi-site, multi-cultural and globally distributed. In short, GSD is software development with dispersed development teams from various geographic locations.

GSD is growing its popularity in many parts of the software industry, for which companies typically adopt nearshoring or offshoring approaches. Nearshoring is the process of moving software development into countries that are in greater proximity to their respective headquarters in terms of time zones, language, and culture (Taplin, 2020). Offshore software development, on the other hand, is distributed across larger distances in geography, time, and culture.

In the US it has been estimated that the value of offshore software development has increased 25-fold over the past 10 years, and furthermore one-quarter of the US spending on application development, integration and management services is estimated to go offshore (Conchúir *et al.* 2009). One of the most common approaches of offshore software development is to use a 'follow-the-sun' approach, where the development is distributed globally in teams between time zones and locations, giving the potential to work on development for 24 hours a day (Taplin, 2020). There are several thought benefits that arise from adopting GSD practices, the main of which is the reduction of development costs. According to Conchúir *et al.* (2006), GSD can lead to faster development cycles and product delivery time as companies can adhere to the 'follow-the-sun' approach, utilizing time differences of time zones. As such, in theory, software can be developed around-the-clock, leading to faster product delivery time and saves in costs. Optimally, GSD allows for businesses to gather a larger pool of talent and skill across borders and save on labor costs.

Furthermore, GSD could optimally allow for closer proximity to customers and markets.

Despite the theoretical advantages of GSD, it contains important risks and challenges that companies need to be aware of. Being geographically distributed across large distances, globally dispersed teams and people across diverse regions face challenges related to temporal and cultural differences. These risks add to challenges in GSD, especially with respect to the coordination of work (Holmström *et al.*, 2006). These problems often result in failing to complete projects on time and running over the budget. A task distribution approach is particularly effective for reducing risks in coordination by minimizing the challenges related to temporal, geographical and socio-cultural distances. Efficient task distribution can help allocate work effectively and solve the challenges related to GSD.

As businesses are increasingly shifting to a more remote yet global way of working, it is essential that we fully understand the challenges that globalization brings to software development, and, more importantly, how to tackle these challenges. What is needed is further research in the challenges related to coordination in GSD, and, furthermore, how effective task distribution models can help to mitigate the challenges of GSD.

1.1 Aim of the Thesis and Research Questions

This thesis starts with a systematic literature review of GSD practices and the challenges related to them. Coordination, communication, and control are identified as major challenges in GSD, posed by its 'global nature'. This is of course due to the larger distances that arise out of the geographical distribution of software development and the dispersed teams and people therein, which are characterized as temporal, geographical, and social-cultural distances. The aim of this thesis is to analyze task distribution models that can mitigate the challenges in coordination and communication in GSD. *Figure 1* shows the conceptual framework which is linked to the structure of the thesis.

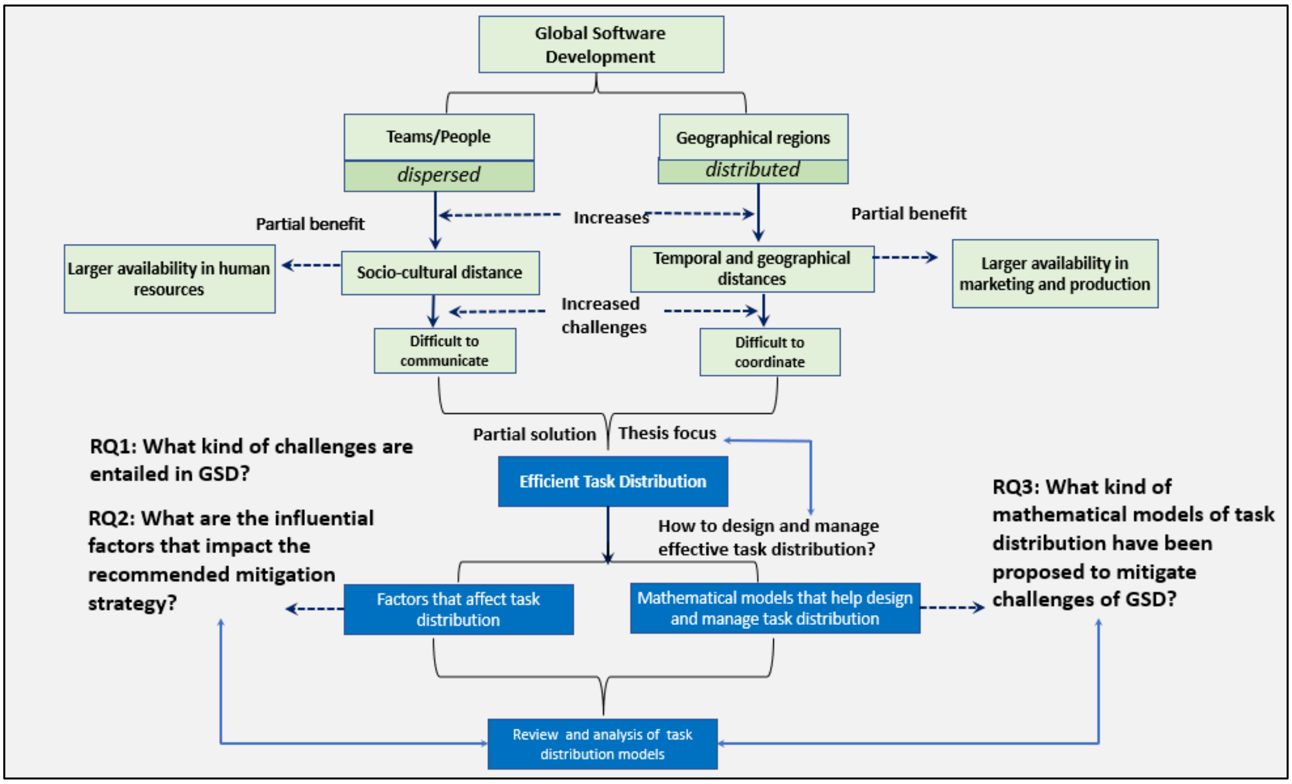

Figure 1. *The conceptual framework of the thesis.*

As such, this thesis aims to understand what type of risk mitigation strategies can be taken to solve GSD challenges. While the literature shows that many mitigation options can help address GSD-related challenges, no single and universal one works for all challenges. Despite this, task distribution models present an effective mitigation approach.

In order to fully understand task distribution models, it is essential to identify (i) *what kind of challenges are entailed in GSD* (RQ1), (ii) *what are the influential factors that impact the recommended mitigation strategy* (RQ2), and (iii) *what kind of mathematical models of task distribution have been proposed to mitigate challenges of GSD?* (RQ3).

As such, the thesis reviews mitigation strategies for solving GSD-related challenges, most importantly task distribution models, focusing especially on mathematical models to ameliorate task distribution in GSD. The research questions are formed based on the conceptual framework shown in *Figure 1*.

1.2 Scope of the Thesis

This thesis especially focuses on reviewing mathematical models intended for efficient task distribution in GSD based on the conceptual framework shown in *Figure 1*. The task distribution models include, for example, task prioritization, classification, and allocation, with the goal to mitigate challenges of GSD.

The thesis will next turn to a brief discussion on the most well-known benefits of GSD, followed by the exploration of common risks in offshore GSD. Risk mitigation strategies are outlined and task distribution models are identified as one of the most effective mitigation strategies (*Section 2: Background*). Then, the methodology is described (*Section 3: Methodology*), after which factors affecting task distribution and mathematical models that help design and manage task distribution in GSD are reviewed (*Section 4: Results*). Finally, the findings are discussed along with the limitations of the study (*Section 5: Discussion*), and a conclusion and potential solutions for the future of work in GSD are provided (*Section 6: Conclusion and the Future of Work*).

2 Background

Already from the 1960s IBM along with other influential IT-companies have increasingly adopted a global software engineering approach (Herbsleb & Mockus, 2003). The amplified wave of globalization since the 1990s has sprung a wave of increased competition in the software industry. However, with a restricted amount of qualified resources and a growing pressure to shorten the time to market, many businesses have now expanded into global markets with an interest to acquire partners and set up software development sites across national borders (*ibid.*). This has led to a new wave of GSD that is multi-site and multi-cultural. The literature concerning GSD and its challenges is vast, as many scholars have attempted to make sense of the challenges facing GSD and other global spheres of business (Carmel, 1999; Herbsleb & Moitra, 2001; Espinosa *et al.*, 2003; Boland & Fitzgerald, 2004; Cherry & Robillard, 2004; Bolici *et al.,* 2009; Kwan *et al.,* 2011; Wickramaarachchi & Lai, 2014).

2.1 Advantages of Global Software Development

2.1.1 Dispersed Teams and People

Teams and the people therein are arguably the most important resource for software development. In GSD, teams and people from multiple nations are involved to jointly develop software. The teams can optimally be built to provide the right people for the right task, drawing from a global, thus larger, pool of talent. With a global business present in various nations GSD brings further advantages, as local knowledge can help with their respective markets. Additionally, the teams interact with different cultures, work experiences, and perspectives on software development problems. The benefits include better access to a large labour pool of skilled employees who can theoretically work around-the-clock on software projects (Hersbleb *et al.*, 2001).

2.1.2 Geographical Distance

The advantages of geographical distances are thought to be, for example, faster development cycles and product delivery as well as a closer proximity to customers and markets. One of the common approaches in GSD is the 'follow-the-sun' approach. Carmel *et al.* (2010) define the approach as software development following the sun into different time zones, optimally enabling around-the-clock work for example in the US and India, as when the US team finishes its work day, the Indians start. As such the building of software can be, theoretically, continuous. This practice is very appealing to companies as it has a promise of reducing the time to market, and businesses can increase their development efficiency with a 24/7 development model. By locating the company's development closer to the customers' location, GSD improves the ability to respond faster to local customer changes and needs (Hersbleb *et al.*, 2001).

2.1.3 Cost Savings

The advantages of dispersed human resources across the globe give rise to low-cost labour and new markets for GSD. This has been amplified by the internet and fast real-time communication across borders. The difference in salaries can vary significantly with regards to geographical distance despite equivalent skills. In a case study conducted by Holmström *et al.* (2006), it was identified that a base annual salary for a software developer in India was US $15 000, which was a quarter of the salary of an Irish developer, which in turn was half of the salary of a developer in the US.

2.2 Challenges in Global Software Development

Despite its theoretical advantages, GSD poses some important challenges such as coordination and communication due to dispersed teams and people across the globe (Cataldo *et al.,* 2007; Haq *et al.*, 2011). Since GSD software projects are complex systems involving dispersed teams and people (*e.g.* various stakeholders from clients to vendors across the globe) and different geographical regions (*i.e.* temporal and cultural distances influencing each other), coordination and communication of GSD becomes challenging. To preserve the benefits of GSD, more coordination and communication for global connectivity is required, however, this can be difficult to achieve. *Table 1* summarises the key challenges in coordination and communication in GSD. These challenges are caused by the dispersed teams and people across the large geographical distances, known as temporal, geographical and socio-cultural distances. The challenges are described below.

Table 1. *Challenges in the coordination of GSD projects.*

Coordination challenges in GSD	**Herbsleb & Mockus (2003)**	**Holmström *et al.* (2006)**	**Herbsleb (2007)**	**Nguyen-Duc *et al.* (2015)**
Delayed response time		X		X
Lack of 'teamness'		X		X
Decrease in communication	X		X	X
Difficulty in finding relevant experience				X
Reduced level of communication efficiency	X		X	
Gap in shared knowledge			X	
Inconsistency			X	
Misinterpretation of tasks		X		X

2.2.1 Temporal Distance

Temporal distance is the amount of time that separates two or more parties that wish to interact together (Holmström *et al.*, 2006). Temporal distance can be caused by time zone difference or different work shift hours that limit the opportunity for real-time collaboration. For example, one hour time zone differences in Europe can possibly lead to added temporal distance because of different working shifts and hours. In contrast, a European company might have less temporal distance to a late working collaborating company in India.

Temporal distance, which is usually large in GSD, can be challenging when it comes to managing projects carried out across different locations. In particular, the delay in response time is seen as a major factor adding to the issue of everyday communication and coordination (Herbsleb, 2007). Holmström *et al.* (2006) further reveal that temporal distance together with limited overlapping hours between workers lead to a delay in response time, causing people to lose track of their work process. This can cause several problems in coordination of GSD and time-critical work.

2.2.2 Geographical Distance

Geographical distance is defined as the measure of effort needed for one party to visit another and it can be seen as a tool for reducing the need for communication. As such, geographical distance can be best measured as the amount of effort needed to relocate compared to the amount of kilometers (*ibid.*). Ease of relocating is measured by several factors, for example time of travel, the requirement of visas, and the opportunity to direct transportation. Generally, a small geographical distance is more beneficial and has more opportunities for collaborating parties.

2.2.3 Socio-Cultural Distance

Ågerfalk *et al.* (2005) posit that 'socio-cultural distance is a measure of an actor's understanding of another actor's values and normative practices.' As such, it is a broad spectrum involving organizational culture, national culture as well as language, politics, motivations, and work ethics on an individual level (Holmström *et al.*, 2006). The most common challenges seem to be related to language and interpretation, as language problems were the primary driver for misunderstandings.

Quick, spontaneous meetings and interactions at work are important for maintaining collaborative relationships with co-workers (Heerwagen, *et al.*, 2004). Informal conversation allows for team members to develop working relationships and acquire an ameliorated flow of information about changes in current projects. The need for informal conversation in GSD is extensive, yet in practice there is far less frequent communication in distributed development teams, with the employees finding it far more difficult to identify distant colleagues and effectively communicate with them (Herbsleb & Mockus, 2003).

In the study conducted by Holmström *et al.* (2006), a common challenge identified among the interviewed was the lack of 'teamness' among the distributed developers. The perks of meeting face-to-face include the forming of social networks that in turn generate trust, respect and commitment in the long run, supporting the development in GSD.

2.3 Task Distribution as a Major Mitigation Strategy

Although many strategies for mitigating GSD challenges have been proposed, for example risk assessment and management of GSD projects (Cichocki & Maccari, 2008), it is effective task distribution that has been considered as the main option in the literature (Carmel & Agarwal, 2001).

Influential factors affecting GSD range from expertise, site characteristics and task site dependency to communication, coordination, and process ownership, with the former three seen as most important by GSD practitioners (Imtiaz & Ikram, 2017). Task distribution models with efficient task allocation that accounts for various influential factors can be seen as important in managing GSD-related challenges.

3 Methodology

A critical literature review with qualitative analysis was chosen as the primary research method in order to gain an understanding of the nature of GSD as well as to identify the research gaps and problems therein.

Consequently, the focus of the research was on task distribution models in GSD. The reviewed papers were collected and searched from commonly known scholarly databases; Elsevier Scopus, IEEE Xplore Digital Library, Google Scholar, and ScienceDirect. The scholarly literature chosen for this thesis consisted of academic papers, journals, conference papers, research case studies and academic reports. The research process included finding the appropriate scholarly literature using search keywords such as 'global software development'; global software development challenges'; 'global software development coordination challenges'; 'coordination in global software engineering' and 'task distribution in global software development'.

In total, 150 articles were collected of which 41 were chosen. A determining factor of the chosen papers was if the temporal, geographical and socio-cultural distances and challenges in GSD were explored in it. This was seen as a significant factor since it laid the foundation to understand the aim and goal of the chosen models, that is, minimizing challenges in GSD by alleviating the challenges related to the above-mentioned distances.

The selected papers for analyzing the factors that contributed to task distribution were cross-referenced amongst other papers of the same topic. After reading several papers on task distribution factors, a clear common message was identified, as many of the papers argued for the most common views supported by many scholars - these are presented in the *Results* section.

3.1 Selected Papers

The selection process of the reviewed task distribution models was structured in three stages: (i) piloting stage, (ii) initial selection, and (iii) final selection. The first stage included the selection of digital publication databases mentioned above and running the keywords on each base. The second stage consisted of discarding and selecting preliminary academic papers and models. This stage involved reading the title, the keywords, the abstract as well as the conclusion of the paper as compared to the research questions and the scope of the thesis. The inclusion and exclusion criteria were also taken into consideration. Subsequently, a list of potential candidate papers was chosen. Then, the candidate publications were read and analyzed, again as compared to the scope and research questions of the thesis. Finally, the results were formed. After the final stage a total of 9 papers and 9 models were chosen. An important factor affecting the selection of the 9 papers was that the they had to be referenced and known in the field.

In addition to selecting and analyzing the publications in the final stage, the final stage also introduced new potential papers on task distribution models through exploring the references of the initial chosen and reviewed papers.

3.2 Inclusion and Exclusion Criteria

The dimension for the scope of the search was defined as limiting it to English publications. English was chosen as the language as it is the most common academic language and the adopted by relevant publishers. The inclusion criteria were defined as publications that match the scope as well as the aim of the thesis, and by being applicable to the chosen research questions. The exclusion criteria were defined as (i) the study did not match inclusion criteria, (ii) it was unable to access for free, and (iii) the study was not available in English.

3.3 Analysis of Papers

This thesis began its research by reading business and technology articles focusing on the topic of globalization and businesses shifting to globally distributed process systems. In order to understand the businesses' motivation for opting for GSD, it was essential to research the theoretical benefits of GSD. After understanding the common benefits and achieving a balanced understanding of GSD the thesis explored some common challenges that are introduced by GSD. By exploring the common challenges, it opened a broader understanding of GSD that involved the connection of temporal, geographical and socio-cultural distances in relation to the challenges.

The established connection of the temporal, geographical and socio-cultural distances in GSD were used as key features for identifying the challenges in coordination and communication in GSD. The challenges were linked to the main mitigation strategies presented in *Figure 1*. To understand and review the current state of task distribution models for GSD, it was essential to identify the factors that contributed to task distribution choices in GSD projects, which are discussed in the conclusion.

4 Results

In GSD the activity of task distribution is much more complicated than expected due to the lack of communication and coordination stemming from temporal, geographical and socio-cultural distances. Furthermore, distributed sites depend on each other because of component, resource, temporal and phase-based dependencies. These dependencies have various coordination needs, rendering them a significant factor on task distribution. The absence of communication and coordination due to above-mentioned distances have been reported by several studies on GSD (Mahmood *et al.*, 2015).

Task distribution decisions directly affect coordination and communication, and when efficient they can minimize the risks of GSD (*ibid.*). Therefore, it is essential to understand the methods that are used to optimize task distribution in GSD. As mentioned above, risks of GSD become evident through temporal, geographical and socio-cultural distances. As such, task distribution for sites, teams and people should reduce the challenges related to the temporal, geographical and socio-cultural distances (Carmel & Agarwal, 2001). Therefore, the literature suggests that the teams and people separated by the said distances should invest in minimizing the distances by putting more effort into team-building, knowledge sharing, communication and coordination as well as the development of module architecture activities (*ibid.*). These are essentially task distribution issues which will be described in more detail shortly.

4.1 Factors Affecting Task Distribution

Task distribution is a complex, multi-layered, and hierarchical process embracing many technical concepts, such as product architecture. Moreover, there are organizational factors stemming from the challenges in geographical, temporal and socio-cultural distances, that influence decisions for task distribution. Mahmood *et al.* (2015; 2017) argue that site technical expertise is the most influencing factor as development sites are distributed in multiple geographical locations, further arguing that choosing a site with the suitable domain expertise is vital for a successful GSD project. Time zone differences are often analyzed and thought to minimize the production time of GSD. Furthermore, Mahmood *et al.* (2015; 2017) argue that resource costs are also an important factor, as businesses often try to assign tasks into geographical regions with low labor costs.

Many influential factors affect task distribution goals in GSD. The literature reviews conducted by Mahmood *et al.* (2015; 2017) summarize and outline the most common factors that influence task distribution decisions in GSD. *Table 2* shows the summary.

Table 2. *Factors that influence task distribution (*Mahmood *et al., 2015; 2017).*

Factors	**Goals**
Site technical expertise	To accomplish tasks. Experience and knowledge management is essential.
Time zone difference	To distribute tasks in advantage typically with ‘follow-the-sun’ approach or ‘time zone-bands’.
Resource cost	To assign work units into sites with lower labor costs.
Task dependency	To support effective interdependencies in order to benefit from GSD projects and decrease development time.
Vendor reliability and maturity	To increase team performance in GSD.
Task size	To minimize production time.
Product architecture	To minimize production time.

4.2 Mathematical Models of Task Distribution in Global Software Development

As described previously, task distribution in GSD is highly complex, involving interrelated influential factors, and as such there are no straightforward guidelines or simple shortcuts for task distribution decisions. Therefore, applied mathematical models are commonly employed to help design and manage task distribution decisions for GSD projects. This section discusses the existing mathematical models that can be used to design and optimize task distribution in GSD.

Table 3. *Task distribution strategy (Filho et al., 2019).*

Stages	Activities	Strategy
(1) Context Characterisation	(1.1) Definition of tasks types (1.2) Identification of influencing factors (1.3) Criteria and groups definition (1.4) Characterisation of influencing factors	Performed once for all participating companies. The data will be shared between companies.
(2) Ranking of the influencing factors	(2.1) Classification of influencing factors into preference groups (2.2) Rank ordering of influencing factors	
(3) Ranking of the executing units	(3.1) Definition and characterisation of offices (3.2) Classification of offices into preference groups (3.3) Rank ordering of offices	Individual per company.

The multi-criteria model (Filho *et al.*, 2019) is proposed as a decision model, applying Verbal Decision Analysis (VDA) to decide which teams should be allocated a given task in the distributed scenario of GSD. VDA is performed to classify and rank the influencing factors and executing units. The approach consists of a workflow accommodating responsible actors and descriptions of the activities. The approach was applied to five companies in the study. Task assignment recommendations are presented in groups for each company, according to the task type, *i.e.*, architecture, requirements, coding, and testing, ordered by the most to the least preferable office. The results show that the approach is adaptable, flexible, as well as easy to understand and use. *Table 3* shows the application result of the approach.

As for the Bayesian network multi-criteria model, Lamersdorf *et al.* (2009) propose Bayesian network-based decision support for task distribution in GSD. Bayesian networks are used to represent the cost functions of the distributed systems algorithm, that is both the cost of executing a task at a site and the cost of transmitting data between sites.

In another study, Lamersdorf *et al.* (2010) argue that a customizable multi-criteria model is best suited for task distribution. The model is able to take into account various (customizable) influencing factors and goals as well as to make project-specific assignment suggestions. Bayesian networks are used to account for uncertainties. Wickramaarachchi & Lai (2013) also argue for a multi-criteria modeling framework for task distribution, accounting for different dependencies in task distribution in GSD, such as work dependency, site dependency, site specific and work specific characteristics, *etc*. Furthermore, Almeida *et al.* (2011) present a multi-criteria decision model for planning and finetuning globally distributed projects in GSD, aimed at enhancing communication and reducing coordination and control overhead.

Alsri *et al.* (2014), on the other hand, propose a model based on neural networks by applying Euclidean distance, which identifies a suitable site for a given task and then finds nearest neighbours to fit the selected sites. The model provides a list of potential sites for the given tasks in order to support the selection of appropriate GSD sites. Conversely, Gupta & Gupta (2021) propose a decentralized, blockchain-oriented, and transparent task distribution framework to improve quality of task allocation processes in GSD in terms of efficient collaboration, good knowledge vaporization, and documentation.

Madachy (2007) proposes a simple parametric cost estimation model for task distribution by phases in an excel format, similarly to Boehm (2000). This is called the statistical analysis model. Sooraj & Mohaptra (2008), on the other hand, opt for an analytical model, suggesting an index model that can be used to evaluate different task distribution alternatives with respect to coordination overhead.

The indices include an indices-coupling index, need index, effectiveness index, and a time zone index, which are created based on a survey, a discrete-event simulation model as well as mathematical equations. *Table 4* summarizes the different modelling methods used for task distribution in GSD.

Table 4. *Summary of the modelling methods used for task distribution.*

Models/Methods	References	Description
Multi-criteria model *Team suggestion for each task*	Filho *et al.* (2019)	Developing a decision model for Verbal Decision Analyis (VDA) to decide which team should be allocated each task in the distributed scenatio of GSD.
Multi-criteria model *Bayesian networks for cost reduction of tasks*	Lamersdorf *et al.* (2009)	Proposing decision support for task distribution in GSD regarding the cost of executing a task at a site and the cost of transmitting data between sites.
Multi-criteria model *Customizable influencing factors and goals*	Lamersdorf *et al.* (2010)	Providing a decision model to take into account various (customizable) influencing factors and goals. Can make project-specific assignment suggestions.
Multi-criteria model *Framework for task distribution*	Wickramaarachchi & Lai (2013)	A framework for different dependencies in task distribution, such as work dependency as well as site specific and work specific characteristics.
Multi-criteria model *Planning and finetuning*	Almeida *et al.* (2011)	Enhances communication and reduces coordination and control overhead.
Neural networks *Suitable sites for specific tasks*	Asri *et al.* (2014)	Provides a list for potential sites for given tasks.
Blockchain technology-based model *Improving the quality of the task distribution process*	Gupta & Gupta (2021)	Provides a task framework to improve the quality of the task distribution process in GSD terms of efficient collaboration, good knowledge vaporization, and documentation.
Statistical analysis model *Cost estimastion*	Boehm (2000); Madachy (2007)	Provides a simple parametric cost estimation model for task distribution by phases in an excel format for GSD.
Analytical model *Index for evaluating tasks*	Sooraj & Mohapatra (2008)	Provides indices such as an indices-coupling index, need index, effectiveness index, and a time zone index.

5 Discussion

This thesis conducted a review of task distribution models proposed by scholars to mitigate the challenges related to GSD. The thesis highlighted the challenges and needs for effective task distribution and, after having reviewed proposed mitigation strategies, it argues that multi-criteria models are the most promising. This is because GSD projects have multiple goals and factors influencing task distribution, and as such, goal and factor conflicts can occur. To avoid this, multi-criteria models are applied to take into account these various goals and factors.

The thesis has shown that there is a growing need for task distribution models in GSD, and we are still far from fully understanding their potential and achieving effective application, for example by taking into consideration the complexity of the factors influencing task distribution choices. Although GSD is a subject that has been studied for some time now, few concrete works on task distribution models exist in the literature. That is, task distribution models have widely been theorized on, yet seldom applied.

5.1 Findings

RQ1: What kind of challenges are entailed in GSD?

Coordination, communication and control as well as challenges from temporal, geographical and socio-cultural distances caused by the dispersed teams and people were identified as the major challenges of GSD .

RQ2: What are the influential factors that impact the recommended mitigation strategy?

All reviewed models have considered several factors that impact task distribution decisions. Most common factors ranged from personal factors such as technical experience and vendors' reliability to geographical factors, such as time zone differences. Task dependencies and product architecture were also mentioned in several of the studies. *Table 1* highlights the factors that impact task distribution models, and, as such, should guide them.

RQ3: *What kind of mathematical models of task distribution have been proposed to mitigate challenges of GSD?*

Many mathematical models have been used to design and manage task distribution in GSD. Nine different models were identified and discussed in the thesis. Among the reviewed models, five were multi-criteria models, one was a neural network model based on machine learning to suggest potential sites for certain tasks, one model was a statistical cost analysis for different phases of task distribution and finally one model was an analytical model for indices in task evaluation. Multi-criteria models were the most adopted model, as is argued in the conclusion.

5.2 Limitations

Due to the space and time constraint, the review is not exhaustive. In addition, the chosen publications may not be representative. The selected sources and the generalization of the findings need further validation.

In addition, the significant number of reviewed task distribution models for solving the challenges in GSD currently exist namely in the scholarly debate and have not been applied in real industrial GSD settings. Even though in the literature the models promise effective results for solving the challenges, the lack of application is still something that should be considered when assessing the validity of the models. The uncertainty of the validity of the models also allows for some space for validity in this thesis.

‘

6 Conclusion and the Future of Work

This thesis has conducted a literature review on GSD and its challenges, analyzed the challenges, and finally suggested task distribution models as the main mitigation strategy. Factors that influence task distribution activity in GSD were described and models that help develop and manage task distribution in GSD were reviewed and analysed.

The following conclusions can be drawn: task distribution in GSD is still too complicated. There are no universal models that can be used for modeling task distribution. In the reviewed papers, a wide range of modeling approaches, such as statistical analysis models, analytical models, neural networks, Bayesian networks and multi-criteria decision models were adopted in designing and managing tasks in GSD. Despite the diversity of the models, multi-criteria models were employed the most among the reviewed publications. Multi-criteria task-distribution models were shown to solve the challenges in coordination and communication in GSD which were the most common reasons for a project failure. This was because the models took into account the various actors and factors that influenced task distribution decisions. While most of the mathematical task distribution models have not been applied to real-life scenarios, the research still shows evidence of a potential benefit of the mathematical models – namely multi-criteria models – on solving challenges in GSD.

Based on this, the thesis suggests for businesses to apply such multi-criteria models to support their task distribution choices when engaging in GSD. Multi-criteria task distribution models are recommended since GSD projects typically have multiple conflicting objectives and, generally, multi-criteria models take these best into account due to the various criteria. As such, it holds the most promise of the various approaches on task distribution. The expected result for businesses adopting multi-criteria task distribution models is a more systematic approach towards the distribution of tasks, reducing the subjectivity of allocation decisions. Ideally this would minimize the challenges in GSD and ensure a successful GSD project delivery.

For the future of work, this thesis proposes further research on the application of multi-criteria task distribution models in real GSD scenarios. With the growing relevance of GSD in software development, understanding the factors that contribute to successful project delivery is essential. Further research in the field of task distribution models will continue to unravel the potential in them. In the long term, researching the applicability of theoretical multi-criteria task distribution models in GSD could give rise to enhanced models, as challenges that are evident in real-life scenarios only can be overcome. Perhaps with further research applied to real industry settings, task distribution models could enhance task allocation in non-GSD contexts as well.

References

Almeida, L. H., Albuquerque, A. B. & Pinheiro, P. R., 'A multi-criteria model for planning and fine-tuning distributed scrum projects', *2011 6th IEEE International Conference on Global Software Engineering*, 2011, pp. 75-83.

Alsri, A., Almuhammadi, S. & Mahmood, S., 'A model for work distribution in global software development based on machine learning techniques', *2014 Science and Information Conference*, 2014, pp. 399-403.

Boehm, B. (2000) *Software Cost Estimation with COCOMO II*. New Jersey: Prentice Hall.

Boland, D. & Fitzgerald, B., 'Transitioning from a co-located to a globally-distributed software development team: a case study and analog devices', *Third International Workshop on Global Software Development*, 2004, pp. 4-7.

Bolici, F., Howison, J. & Crowston, K., 'Coordination without discussion? Socio-technical congruence and stigmergy in free and open source software projects', *Second International Workshop on Socio-Technical Congruence*, 2009, pp. 1-9.

Carmel, E. (1999) *Global Software Teams: Collaborating Across Borders and Time Zones*. New Jersey: Prentice Hall.

Carmel, E. & Agarwal, R. (2001) 'Tactical approaches for alleviating distance in global software development', *IEEE Software*, (18) 2, pp. 22-29.

Carmel, E., Dubinsky, Y. & Espinosa, J. A. (2010) 'Follow the sun workflow in global software development', *Journal of Management Information Systems*, (27) 1, pp. 17-38.

Cataldo, M., Bass, M., Herbsleb, J. D. & Bass, L., 'On coordination mechanisms in global software development', *International Conference on Global Software Engineering (ICGSE 2007),* 2007 pp. 71-80.

Cherry, S. & Robillard, P. N., 'Communication problems in global software development: spotlight on a new field of investigation', *26th International Conference on Software Engineering,* 2004, pp. 48-52.

Cichocki, P. & Maccari, A. (2008) 'Empirical analysis of distributed software development project', *Balancing Agilicty and Formalism in Software Engineering*, (5082), pp. 169-181.

Conchúir, E. Ó., Holmström, H., Ågerfalk, P. J. & Fitzgerald, B., 'Exploring the assumed benefits of global software development', *IEEE International Conference on Global Software Engineering (ICGSE'06),* 2006, pp. 159-168.

Conchúir, E. Ó., Ågerfalk, P. J., Olsson, H. H. & Fitzgerald, B. (2009) ‘Global software development: where are the benefits’, *Communications of the ACM*, (52) 8, pp. 127–131.

Espinosa, J. A., Cummings, J. N., Wilson, J. M. & Brandi, M. P. (2003) ‘Team boundary issues across multiple global firms’, *Journal of Management Information Systems*, (19) 4, pp. 157-190.

Filho, M. S., Pinheiro, P. R., Albuquerque, A. B., Simao, P. S., Azevedo, R. S. & Nunes, L. C. (2019) ‘A multicriteria approach to support task allocation in projects of distributed software development’, *Complexity*, (2019), pp. 1-22.

Gupta, C. & Gupta, V. (2021) ‘A decentralized framework for managing task allocation in distributed software engineering’, *Applied Sciences,* (11) 22, pp. 1-10.

Haq, S., Raza, M., Zia, A. & Khan, M. (2011) ‘Issues in global software development: a critical review’, *Journal of Software Engineering and Applications*, (4) 10, pp. 590-595.

Heerwagen, J. H., Kampschroer, K., Powell, K. M. & Loftness, V. (2004) ‘Collaborative knowledge work environments’, *Building Research and Information*, (32) 6, pp. 510-528.

Herbsleb, J. D. & Moitra, D. (2001) ‘Global software development’, *IEEE Software*, (18) 2, pp. 16-20.

Herbsleb, J. D. & Mockus, A. (2003) ‘An empirical study of speed and communication in globally distributed software development’, *Software Engineering, IEEE Transactions*, (29) 6, pp. 481-494.

Herbsleb, J. D. ‘Global software engineering: the future of socio-technical coordination’, *Future of Software Engineering (FOSE’07)*, 2007, pp. 188-198.

Holmström, H., Fitzgerald, B., Ågerfalk, P. J. & Conchúir, E. Ó. (2006) ‘Agile practices reduce distance in global software development’, *Information Systems Management*, (23) 3, pp. 7-18.

Holmström, H., Ågerfalk, P. J. & Fitzgerald, B., ‘Global software development challenges: a case study on temporal, geographical and socio-cultural distance’, *2006 IEEE International Conference on Global Software Engineering (ICGSE’06)*, 2006, pp. 3-11.

Imtiaz, S. & Ikram, N. (2016) ‘Dynamics of task allocation in global software development’, *Journal of Software: Evolution and Process*, (29) 1, pp. 1-17.

Kwan, I., Schroter, A. Damian, D. (2011) ‘Does socio-technical congruence have an effect on software build success? A study of coordination in a software project’, *IEEE Transactions on Software Engineering,* (37) 3, pp. 307-324.

Lamersdorf, A., Münch, J. & Rombach, D. (2009) ‘A decision model for supporting task allocation processes in global software development’, in Bomarius, F., Oivo, M., Jaring, P., Abrahamsson, P. (eds) *Product-Focused Software Process Improvement*. Berlin: Springer.

Lamersdorf, A. & Münch, J. (2010) ‘A multi-criteria distribution model for global software development projects’, *J Braz Comput Soc,* (16), pp. 97-115.

Madachy, R. (2007) ‘Distributed global development parametric cost modeling’, in Wang, Q., Pfahl, D., Raffo, D. M. (eds.) *Software Process Dynamics and Agility*. Berlin: Springer.

Mahmood, S., Anwer, S., Niazi, M., Alshayeb, M., & Richardson, I., ‘Identifying the task factors that influence task allocation in global software development: preliminary results’, *Proceedings of the 19th International Conference on Evaluation and Assessment in Software Engineering (EASE'15)*, 2015, pp. 1-6.

Mahmood, S., Anwer, S., Niazi, M., Alshayeb, M., & Richardson, I. (2017) ‘Key factors that influence task allocation in global software development’, *Information and Software Technology*, (91), pp. 102-122.

Nguyen-Duc, A., Cruzes, D. & Conradi, R. (2015) ‘The impact of global dispersion on coordination, team performance and software quality – a systematic literature review’, *Information and Software Technology*, (57), pp. 277-294.

Olson, J. & Olson, G. (2000) ‘Distance Matters’, *Human-Computer Interaction*, (15) 2, pp. 139-178.

Sooraj, P. & Mohapatra, P., ‘Developing an inter-site coordination index for global software development’, *2008 IEEE International Conference on Global Software Engingeering*, 2008, pp. 119–128.

Taplin, S. (2020) ‘Offshore versus Nearshore Software Development’. Forbes, 10 December. Available at: https://www.forbes.com/sites/forbestechcouncil/2020/12/10/offshore-versusnearshore-software-development/?sh=1b10fb9275bf [Last accessed 17 April 2022].

Wickramaarachchi, D. & Lai, R., ‘A method for work distribution in global software development’, *Proceedings of the 2013 3rd IEEE International Advance Computing Conference (IACC)*, 2013, pp. 1443.1448.

Wickramaarachchi, D. & Lai, R., ‘Software modularization in global software development’, *2014 International Conference on Data and Software Engineering*, 2014 *(ICODSE),* pp. 1-6.

Ågerfalk, P. J., Fitzgerald, B., Holmström, H., Lings, B., Lundell, B. & Conchúir, E. Ó., 'A framework for considering opportunities and threats in distributed software development', *Proceedings of the International Workshop on Distributed Software Development,* 2005*,* pp. 47-61.