AJIT-e: Online Academic Journal of Information Technology
2018- Informatics and Communication Technologies Special Issue/İletişim ve Bilişim Teknolojileri Özel Sayı –Cilt/Vol: 9-Sayı/Num: 35
DOI: 10.5824/1309-1581.2018.5.011.x
http://www.ajit-e.org/?menu=pages&p=details_of_article&id=375

Received : 26.09.2018 Editorial Process Begin: 20.10.2018 Published: 11.12.2018

A Literature Review on Challenges in Distributed Software Development

Murat Doğuş KAHYA, *Siemens A.S., Istanbul/Turkey, murat.kahya@siemens.com*
Çağla ŞENELER, *Yeditepe University, Istanbul/Turkey, cagla.seneler@yeditepe.edu.tr*

ABSTRACT *Distributed software development (DSD) becomes a trend for most of the global companies in the last decade. According to the effects of distributed development, new challenges exist that are not observed at collocated software development. The Software Engineering community investigated the challenges more than a decade that was mostly consisted of communication, coordination and control processes. However, most of the latest studies were empirical and did not cover all challenges in the industry. In this paper, GSD challenges are reported as a systematic literature review with temporal, geographical and socio-cultural dimensions. This research aims to create a guideline of challenges that may exist in the GSD projects years between 2007 and 2017. The discussion of implications for practitioners and future research are mentioned at the end of the report.*

Keywords*: Global software development, Distributed software development, Challenges, Problems, Systematic literature review*

Literatür Taraması: Dağınık Yazılım Geliştirmede Yaşanan Zorluklar

ÖZ *Küresel şirketler tarafında son on yılda farklı lokasyonlarda yazılım geliştirme popüler olmaya başladı.Farklı lokasyonlarda geliştirme ile birlikte, aynı lokasyonda yazılım geliştirmede yaşanmayan yeni zorluklar ortaya çıkmaya başladı.Bununla birlikte yazılım geliştirme toplulukları iletişim, koordinasyon ve kontrol süreçlerinin yaşadığı zorlukları son yıllarda araştırıyor. Fakat son yıllarda yayınlanan çalışmalar çoğunlukla bir kaç şirketi göz önünde bulunduran deneysel ve vaka çalışmalarını içermektedir. Bu çalışmada 2007 ile 2017 yılları arasındaki deneysel çalışmalar derlenerek zamansal, coğrafik ve sosyo-kültürel açılarından yaşanan zorluklar literatür incelemesi olarak sunulacaktır. Makale kapsamında 607 çalışma incelenerek bulunan sonuçlar araştırmacılar ile paylaşılmaktadır.Sonraki araştırmalar ve çıkarımlar raporun en sonunda sunulmaktadır.*

Anahtar Kelimeler: *Küresel yazılım geliştirme, Dağınık yazılım geliştirme, Zorluklar, Problemler, Sistematik literatür araştırması*

Introduction

It is known that globalization is applied by most of the companies to expand their business all over the world. Within the effects of globalization, software development practices and workflows are also impressed. The traditional way of software development; product management and development teams worked at the same location. Currently, the situation for most of the teams is changed as working at different geographical locations as the matter of the fact members of development

team may be at different locations in the current environment. As a result of globalization of companies, integration and transformation of DSD is started. In the literature, DSD is known as Global Software Development (GSD) and also Global Software Engineering (GSE) that is defined as software development teams in different geographical locations Karolak (1998). Software development is a complex process that needs to have good collaboration among team members; therefore, DSD creates a more complex and challenging environment. According to Karolak, DSD projects are considered as more complex to conduct than the most collocated in-house project (Jabangwe & Nurdiani, 2010).

As well as, some companies try to take advantages of having development teams at different geographical locations. These advantages can be enhancing time-to-market, access to cheap skilled labor, gaining competitive advantages and availability of in-house employees to prevent or overcoming possible problems comparing to rivals who do not apply DSD (Moe & Smite, 2008).

In addition to this, adoption of distributed development comes with benefits such as delivery time to market, being closer to potential clients and also satisfying the local market demands for international companies. In the meantime, companies get a chance to collect different perspectives from various geographical locations in the world that may create a good sense of usability, local norms, and scalability of the product. Moreover, discovering new ideas is easier than before by the mixture of developers with different cultural backgrounds and geographical locations that can be regarded as another favorable impact of DSD.

Ideally, software development team members are working at the same location. Thus, developers have face-to-face meetings, same culture, and organization that enhance the collaboration of team members and also makes easier to control. Within the existence of DSD, team members work at distributed locations that cause challenges such as different time-zones, communication difficulties, social and cultural differences. Furthermore, team members need to make up a network of distributed sites due to locating in various isolated locations during software development lifecycle. According to Smite, the team works across geographical, cultural, temporal, organizational and political boundaries to success a common software project, defined as DSD team (Holmstrom, Conchúir, Ågerfalk, & Fitzgerald, 2006). According to Holmstrom, distributed development brings up new dimensions to the software development process that can be mentioned as temporal, geographical and socio-cultural distances (Kitchenham & Charters, 2007).

DOI: 10.5824/1309-1581.2018.5.011.x

The objective of this paper is to present challenges that associate with DSD by describing key elements that may create risks for project processes such as communication, coordination, and control in terms of the current industry environment. In this context, challenges from the empirical studies are categorized as the temporal, geographical and socio-cultural distance in the results section.

Research Questions

A systematic review is a form of secondary study that reviews all primary studies that are relevant to research question for integrating research paper (Kitchenham & Charters, 2007). In order to give the definition of primary study, "an empirical study investigating a specific research question" (Kitchenham & Charters, 2007).

In this research, SLR has adopted the guideline that was written by Kitchenham. The guideline was followed to facilitate this research by identifying, evaluating and interpreting studies that were mostly including a review of empirical studies (Kitchenham & Charters, 2007).

The main reasons to conduct SLR for this research;

- To identify challenges of GSD in the literature
- To provide a background for new researches from the collected challenges those were systematically analyzed with the help of empirical studies in the literature.

Regarding the enhanced understanding of the challenges reported that are consisted of empirical studies, following research questions are considered;

- RQ.1 Which GSD challenges are reported in literature related to temporal distance?
- RQ.2 Which GSD challenges are reported in literature related to geographical distance?
- RQ.3 Which GSD challenges are reported in literature related to socio-cultural distance?

Research Method

This research is conducted to identify challenges from the empirical studies that were subjected to distributed software development in the literature. This review consists of two main strategies;

- The review of studies that were mentioned in the selected literature review to combine and eliminate related studies.
- The search for primary studies in the literature to find challenges that were not reported in the selected literature review

In this section, the study selection, data extraction processes and threats to the validity of are presented in detail.

Data Retrieval

In this section, the secondary study selection process is presented in detail. In order to find out systematic reviews in the literature, a search conducted with four significant repositories related to research topic; IEEE Xplore Digital Library, Science Direct, Sprinter and ACM Digital Library. Following search string is used to find the information;

({Global software development} OR {Distributed software development} OR {Global software engineering} OR {Distributed software engineering}) AND ({systematic review} OR {systematic literature review})

35 papers were found from this search string, 13 of them were duplicated at four repositories. Two of them were related to adaptation of social network in GSD. Six papers were mostly focused on effort estimation and performance issues for scrum teams in GSD which were not related to this research.

There are two tertiary studies that were focused on the systematic reviews to review and rate them by considering the quality and citation count. With the help of the tertiary studies, one paper by Rizvi, Bagheri, and Gasevic (Rizvi, Bagheri, and Gasevic, 2015), was selected for a based paper that was focused on reported problems of empirical studies in GSD. Except that, this review was focused years between 2007 and 2012, therefore there is a need for covering latest empirical publications of GSD. However, the availability of resources after 2012 was limited in this paper therefore SLR process was done from scratch years from 2012 to 2017.

DOI: 10.5824/1309-1581.2018.5.011.x

Study Selection

GSD has been a trend in the 21st century with the technological improvements in communication channels that enabled connection between distributed locations (Holmstrom, Conchúir, Ågerfalk, & Fitzgerald, 2006) (Herbsleb & Mockus, 2003). However, communication technology was mostly changed in recent years; people can talk another part of the world anytime wanted. Because of this consideration and for the purpose of the thesis, papers were not found as relevant and not included that are published prior to 2007. Therefore research papers published from 2012 to 2017 were included in this review.

The review of publications is the main focus of the SLR by investigation and analysis of empirical studies according to the industry. Gasevic et al. published a systematic review of empirical studies from 2007 to 2012 (Rizvi, Bagheri, and Gasevic, 2015). Therefore there is a requirement to extend the study conducted from Gasevic et al. by using the similar strategy that focuses on the years between 2012 and 2017 to identify recent challenges in GSD environment.

IEEE Xplore Digital Library, ACM Digital Library, Science Direct, and Springer are used for the search for empirical studies with the help of the following search string;

("Global software development" OR "Global software engineering" OR "Distributed software development" OR "Distributed software engineering") AND ("empirical" OR "experiment" OR "case study")

Moreover, there is a need to narrow down papers that were candidates for the literature review process. Hence, this thesis is including only full-text papers to ensure for studying on the relevant publications. The publication selection criteria and procedures are given in Table 1;

Table 1 - Publication Selection Criteria

Selection of Study	Inclusion Criteria
Based on the search string	Only English Only published papers Publication date after 2012 till 2017
Based on title	No slides, panels, editorials, discussions, summaries of tutorials, duplicates or comments
Based on abstract	Emphasize on empirical studies in relation to GSD
Based on the complete text	An investigation based on GSD challenges Presence of empirical data

In addition to these, all steps are presented in Figure 1 in terms of publication selection criteria.

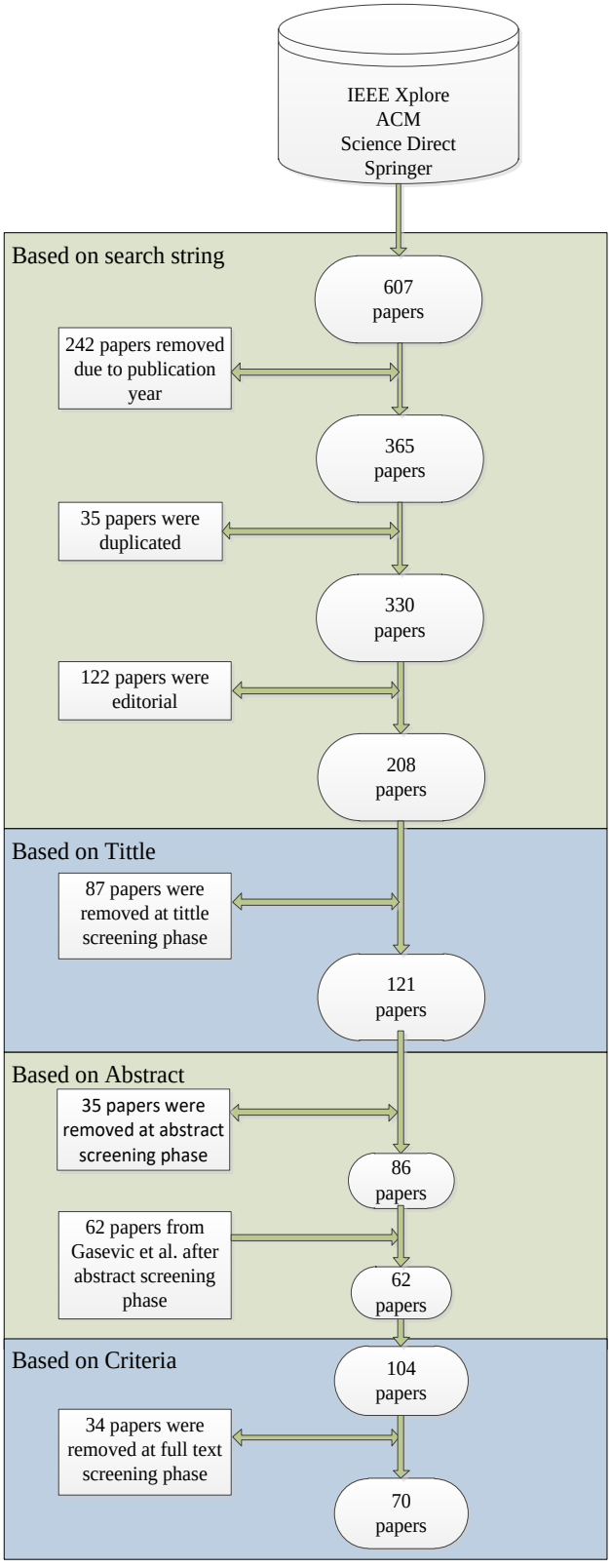

Figure 1 - Systematic Literature Review Steps

DOI: 10.5824/1309-1581.2018.5.011.x

Data Extraction

For each of the 84 selected studies, paper title, date of publication, research methodology, subject of empirical study, software development methodology and challenges were extracted.

Each challenge was categorized with the help of scheme offered by Holmstrom et al (Holmstrom, Conchúir, Ågerfalk, & Fitzgerald, 2006) that was an empirical study on three international software development companies. Challenges are grouped with three main parts as temporal, geographical and socio-cultural distance.

In order to include different perspective to the categorization of the challenges, one more researcher invited to analyze and categorize challenges to the given schema. Face-to-face meetings were organized to discuss categorization of challenges between main author and invited researcher.

Threats to Validity

For the assessment of the validity of this study, two steps of validity are considered in this section. Firstly, reliability threats are considered that is related with the identification of same findings from the other researchers and also replication of the study. This issue is solved with peer review of authors independently from each other to minimize bias. For the second step, investigation for searching papers is started with IEEE Xplore digital library that is known for the collectivity of Software Engineering topics.

However, there is a need for corroborating the list of papers to have complete coverage of the period. Therefore another search was done at ACM, Science Direct and Springer digital libraries. Moreover, found papers might consist from local venues as Asia-Pacific Software Engineering Conference which cannot be generalized for global software development and are removed from the list if the conferences are not indexed in IEEE and ACM database.

Moreover, conference papers were most suitable for this study that consisted as seen in Figure 1;

- 39 papers from conferences
- 31 paper from journals

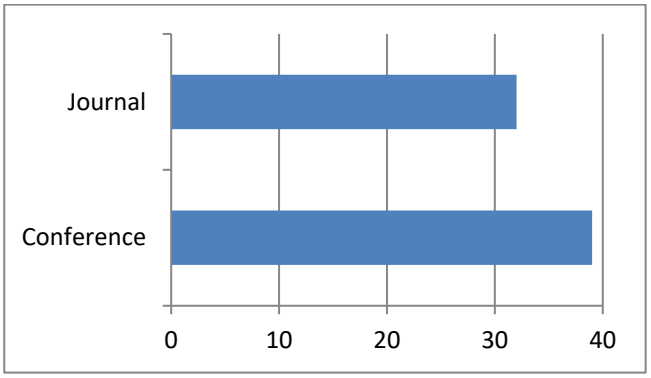

Figure 2 - Distribution of Selected Studies by Type

Results

The results of the systematic literature review are summarized in three categories as proposed from Holmstrom et al (Holmstrom, Conchúir, Ågerfalk, & Fitzgerald, 2006). Only the references of selected studies were included in the tables due to lack of space for discussion at each item in detail.

Temporal Distance

The most effective way for communication is having face-to-face communication which was a basic prototype for communication (Clark, 1996) (Carmel, 1999). Nonetheless, development teams are dispersed in the distributed development which disables the face-to-face communication (Holmstrom, Conchúir, Ågerfalk, & Fitzgerald, 2006).

In addition, dispersion of development sites might be located at different time-zones that make impossible to have synchronous communication with other team members. Temporal distance makes difficult to have effective communication with other locations by the reason of lack of face-to-face and asynchronous communication.

Table 2 - Challenges of Temporal Distance

	Challenges	References
Temporal Distance	*Effective Communication*	R3,R5,R6,R9,R10,R15,R16,R17,R18, R25,R28,R30,R31,R35,R41,R42,R44, R49,R52,R54,R55,R56,R63,R64,R65, R66,R67,R68,R69,R71
	Coordination	R6,R36,R39,R40,R48,R55,R56,R66,R67,R68,R71
	Time Difference	R4,R17,R27,R28,R30,R36,R40,R42,R53,R54,R58,R59, R63,R68,R69
	Delayed communication feedback	R6,R33,R38,R39,R55,R56,R66,R68,R71

DOI: 10.5824/1309-1581.2018.5.011.x

Geographical Distance

The definition of Geographical Distance is a measure for the effort to reach one person to another which can be evaluated as reducing the communication between developers in distributed software development (Holmström, Fitzgerald, Ågerfalk, & Conchúir, 2005). Miles or kilometers are not highly relevant to the geographical distance; mostly it was connected with the ease of relocation. If the direct medium of transportation from one location to another place exists, geographical distance might not be a huge challenge for the GSD projects.

Table 3 - Challenges of Geographical Distance

	Challenges	References
Geographical Distance	*Lack Of Trust*	R7,R8,R13,R20,R33,R34,R48,R49,R50,R56,R60
	Asymmetry in Processes, Policies, and Standards	R3,R5,R10,R16,R20,R24,R26,R37,R38,R50,R55,R62,R68,R70
	Physical Distance/ Geographic dispersion	R9,R11,R19,R40,R43,R49,R52,R56R57, R58,R59,R63,R64,R65,R67, R68,R69
	Organizational Structure	R3,R47,R55,R65
	Product Architecture	R5,R6,R11,R12,R19,R20,R31,R44,R45,R65
	Customer Involvement	R4,R8,R11,R29,R33,R36,R58
	High communication cost	R15,R46,R49,R52,R55,R58,R60, R64
	Poor communication infrastructure	R1,R3,R4,R5,R9,R10,R17,R29,R30,R53, R54, R55,R68
	Critical knowledge in the hands of a few people	R1,R6,R33,R45
	Lack of team cohesiveness	R2,R3,R11,R14,R17,R27,R28,R36, R55,R56
	Limited possibility for informal communicating due to dispersion of sites	R2,R7,R8,R14,R16,R17,R21, R26,R27,R28,R39,R49,R50, R54,R55

Socio-Cultural Distance

The measure of understanding between people from different cultures defined as socio-cultural distance (Holmstrom, Conchúir, Ågerfalk, & Fitzgerald, 2006). The effect of culture is an unrepudiated truth for the given reactions to specific situation or

interaction with other people. On the other hand, the existence of socio-cultural distance among distributed teams become unavoidable if they are using different native language and backgrounded with different cultures.

Table 4 - Challenges of Socio-Cultural Distance

	Challenges	References
Socio-Cultural Distance	*Language*	R4,R17,R18,R42,R43,R45,R48,R49,R56,R58,R61,R64,R67
	Cultural Differences	R1,R12,R16,R18,R27,R30,R40,R42,R43,R48,R58,R59,R61,R64,R66
	Different knowledge levels	R2,R5,R6,R7,R10,R13,R31,R37,R39,R44,R50R52,R57,R59,R60,R61,R62,R68,R70
	Fear of job loss	R61

Conclusion

Global software development becomes vital for the most of the companies that would expand business all over the world. Different nations and cultures expect different types of product from each other that make global software development more significant. According to distributed development, challenges are existing due to the time difference, cultural issues, physical distance etc. In connection with this, Software Engineering community is investigating these challenges to find out mitigation strategies. This paper aimed to create a baseline to facilitate future researchers with current challenges in GSD.

According to the results section of this paper, geographical distance is the most significant issue that may create problems and risks for the GSD projects. Regarding this, most of the challenges were caused by physical distance and effective communication when getting to the root of the problem. Therefore companies may select carefully development teams' time zones that are planned to work together.

DOI: 10.5824/1309-1581.2018.5.011.x

Implications for Practitioners

This study is focused on the challenges and problems in GSD to help practitioners for identification of state-of-art. Practitioners can descend to particulars of specified challenges by checking the references that may create huge time saving for them.

Implications for Future Works

Large amount of challenges are presented in this study to create a state-of-art for the researchers. There were lots of empirical studies that mainly focused on applied solutions from the industry and academy. However solutions and mitigation strategies associated with challenges are missing as a state-of-art. Moreover, validation of reported solutions is a need to make improvements in global software development process.

REFERENCES

1. R. Sangwan, M. Bass, N. Mullick, D. Paulish, and J. Kazmeier, Global Software Development Handbook. Auerbach Publications. 2006.

2. Karolak, D. W. Global Software Development – Managing Virtual Teams and Environments. Los Alamitos, IEEE Computer Society, USA, 1998, pp.159.

3. Jabangwe, R., and Nurdiani, I. Global Software Development Challenges and Mitigation Strategies A Systematic Review and Survey Results. Ronneby, Sweeden: Blekinge Institute of echnology. 2010.

4. Moe N.B., Šmite D. (2007) Understanding Lacking Trust in Global Software Teams: A Multi-case Study. In: Münch J., Abrahamsson P. (eds) Product-Focused Software Process Improvement. PROFES 2007. Lecture Notes in Computer Science, vol 4589.

5. Holmstrom, H., Conchúir, E. Ó., Ågerfalk, P. J., & Fitzgerald, B. (2006). Global Software Development Challenges: A Case Study on Temporal, Geographical and Socio-Cultural Distance. IEEE International Conference on Global Software Engineering (ICGSE'06). IEEE.

6. Kitchenham, B., & Charters, S. (2007). Systematic literature reviews in software engineering - A systematic literature review. Keele, UK: Vol 2.3 EBSE Technical Report, EBSE-2007-01, Software Engineering Group.

7. Rizvi, B., Bagheri, E., & Gasevic, D. (2015). A systematic review of distributed Agile software engineering. Journal of Software: Evolution and Process, 27(10), 723–762. https://doi.org/10.1002/smr.1718

8. Herbsleb, J. and Mockus A., "An empirical study of speed and communication in globally distributed software development," Software Engineering, IEEE Transactions on, vol. 29, 2003, pp. 481-494.

9. Clark, H. (1996). Using Language. Cambridge: Cambridge University Press.

10. Carmel, E. (1999). Global Software Teams: Collaborating Across Borders and Time Zones. Upper Saddle River: Prentice Hall.

11. Ågerfalk, P., Fitzgerald, B., Olsson, H., Lings, B., Lundell, B., & Ó Conchúir, E. A framework for considering opportunities and threats in distributed software development. (2005, January 1).