Identifying Software Cost Attributes of Software Project Management in Global Software Development: An Integrative Framework

Manal El Bajta
Software Project Management Research Team
ENSIAS, Mohammed V University in Rabat
Rabat, Morocco
manal.elbajta@gmail.com

Ali Idri
Software Project Management Research Team
ENSIAS, Mohammed V University in Rabat
Rabat, Morocco
ali.idri@um5.ac.ma

ABSTRACT

The management of global and distributed software projects is a very difficult task further complicated by the emergence of new challenges inherent in stakeholder dispersion. Software cost estimation plays a central role to face challenges in the context of Global Software Development (GSD). The objective of this study is to identify software cost attributes related to GSD context to present an integrative framework encompassing these attributes. Thirty cost attributes were identified using a Systematic Literature Review (SLR) and later compiled into a framework inspired by the Software Engineering Institute (SEI) taxonomy.

CCS CONCEPTS

• **Software and its engineering → Collaboration in software development**;

KEYWORDS

Cost attributes, Software Project Management, Global Software Development, Systematic Review

ACM Reference Format:
Manal El Bajta and Ali Idri. 2020. Identifying Software Cost Attributes of Software Project Management in Global Software Development: An Integrative Framework. In *13th International Conference on Intelligent Systems: Theories and Applications (SITA'20), September 23–24, 2020, Rabat, Morocco.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3419604.3419780

1 INTRODUCTION

Global Software Development (GSD) is a manner of producing software where the life cycle of software development involves stakeholders from remote locations and possibly multiple social backgrounds. This phenomenon began in the early years of this millennium and has evolved to a widely recognized and embraced approach to software development from a newly adopted practice. The growth of offshore software development outsourcing indicates the need to better understand associated issues or problems [1].

SITA'20, September 23–24, 2020, Rabat, Morocco

ACM ISBN 978-1-4503-7733-1/20/09…$15.00
https://doi.org/10.1145/3419604.3419780

Identification of software cost attributes is one of the most important processes of Software Project Management (SPM) and has been widely discussed in literature [2, 9, 23]. Interest in Global Software Development (GSD) is rapidly growing [12], identifying the new challenges associated with this phenomenon has push researchers to investigate more to select relevant list attributes adapting to a specific framework and then compiled using survey researchers and practitioners of software project managers involved in GSD.

This paper provides the investigation of software cost factors used to address the software cost estimation related to the management of GSD projects that are available in literature. The aim of this study is to develop a framework encompassing these attributes using a list of software cost attributes that answer the question of "what" is to be built in the GSD context.

In this paper, we present the results of our study in identifying challenges related to SPM in GSD and categorizing them according to the selected model. The list of software cost attributes has been acquired through a Systematic Literature Review (SLR) of studies focusing on the approaches facing project managers of GSD projects. To do this, we intend to address the following Research Question (RQ):

RQ: What are the software cost attributes of software project management in GSD context?

The paper is structured as follows. Sect. 2 discusses related work. The research methodology is explained in Sect. 3. Results from the SLR are presented and discussed in Sect. 4. Threats to Validity are presented in Sect. 5. Finally, conclusions are presented in Sect. 6.

2 RELATED WORKS

One of the earliest researches in software cost estimation has been carried out by Heemstra [15] prior to the emergence of GSD context. He emphasized the reasons for overruns of cost, and the prerequisites for estimating in the development software life-cycle to lessen software budgets, and avoid software cost estimation challenges. He presented an overview of software cost estimation models with a framework listing the factors which are commonly regarded as important structured in five categories. The emergence of GSD has profoundly changed the way software cost approaches are developed. Some of the factors that are presented in collocated software development have a more critical impact on the outcome of GSD projects while they also face new challenges innate to their distributed setting. This change highlights the need of new tailored approaches to the software cost estimation for GSD projects [3, 13].

Keil et al. [19] presented in 2006 an additional cost drivers of distributed development, and examined the significance of each of these factors as a contributor to the overall cost of a software development project. The paper aims to identify 18 effort-multiplier cost drivers related to SPM for GSD. It classifies drivers into four categories (product, platform, personnel, and project factors) and was inspired from previous study carried out by boehm et al. [4].

In 2006, Smite et al. [33] published an article presenting a framework for overcoming threats in global projects. The research aims to investigate the nature of global risks and build a comprehensive framework presenting global factors and threats that distinguish distributed projects. The data was gathered from literature [10, 32], and analyzed according to principles prescribed by a grounded theory through applying open, axial, and selective coding techniques [35]. The study derived several global factors from the performed taxonomy of software development risks, developed by Software Engineering Institute [8]. The results of this paper show that these frameworks fail to cover all software cost factors presented in literature mainly because the lack of investigation tackled by researchers. In this paper, we aim first to lay out an updated check-list of software cost factors related to managing GSD projects and then to provide an integrative framework for identifying and classifying these factors.

3 METHODOLOGY

An SLR is a means of identifying, evaluating and interpreting all available research relevant to particular RQ [18]. One of the most common reasons for undertaking an SLR aim to provide a framework/background in order to appropriately position new research activities. Protocol of this systematic review has been written based on the recommendations of Kitchenham and Charters [18] to describe the review process, three main points are then extracted:

> 1- Search strategy: Identification of RQs, construction of search string and definition of relevant libraries.
> 2- Study selection: Conduct the review and perform the selection criteria.
> 3- Data extraction and analysis: extraction, monitoring and synthesis.

3.1 Search strategy

This SLR aims to answer the following RQ:

RQ: What are the relevant software cost attributes in GSD context?

Using the PICO method (Population, Intervention, Comparison, Outcome) on the RQ, the search string have been identified:

- Population: Global Software Development
- Intervention: Cost estimation
- Outcome: Attributes

Synonyms for each of search terms are identified based on previous studies. They are presented in Table 1.

The search string used to perform the automatic searches in the digital libraries selected was formulated using the OR logical operator between synonyms and using the AND logical operator between combined terms of the PICO method. The search string used is:

Table 1: Synonyms of the initial search terms

Population	Intervention	Outcome
global software development	cost estimat*	attribute
distributed project	cost plan*	driver
dispersed development	cost predict*	factor
offshor*	cost measur*	
outsourc*	cost calcul*	

(Global software development **OR** distributed project **OR** Dispersed development **OR** outsourc* **OR** Offshor*) **AND** (cost estimat* **OR** cost plan* **OR** cost predict* **OR** cost measur* **OR** cost calcul*) **AND** (attribute **OR** driver **OR** factor)

The extracted papers were identified using the specified search terms, and the searches have been performed in the following sources in January 2018:

- IEEE Xplore digital library [1]
- Association for Computing Machinery (ACM) digital library [2]
- ScienceDirect [3]
- SpringerLink [4]

3.2 Study selection

This search string was applied to the titles, abstracts, and keywords of papers to reduce the search results. Each paper was retrieved by two researchers separately who each identified relevant studies. When there was a disagreement between the two researchers, in a meeting between all the researchers involved in this study, the full text was accessed to reach an agreement. We defined the following inclusion and exclusion criteria based on established SLR guidelines [6]:

Inclusion criteria

- IC1: Papers related to software cost aspect of managing GSD projects.

Exclusion criteria

- EC1: Papers that are not in English.
- EC2: Papers that are published before 2010.

In this SLR, the quality assessment process was not conducted. The main reason behind this is to propose a comprehensive list of software cost attributes related to SPM in GSD context found in academic literature.

3.3 Data extraction

The data extraction strategy was carried out by the main author in January 2018 and reviewed by the remaining authors. The following

[1] http://ieeexplore.ieee.org
[2] http://dl.acm.org
[3] http://sciencedirect.com
[4] http://link.springer.com

data was extracted from each paper by obtaining a set of possible answer to the RQ.

Software Engineering Institute (SEI) taxonomy is developed to support software cost management in different kinds of software development projects, while our aim was to facilitate cost estimation in GSD [29], [28], [21]. The taxonomy organizes software cost estimation threats into three major classes: Product Engineering, Project Personnel, and Development Environment. We used a SEI taxonomy to classify the identified and uncovered global software cost attributes. Boehm et al. [5] defines these three categories as follows:

- Product Engineering: focuses on the nature of product engineering and process, product requirements, and project constraints.
- Project Personnel: focuses on the resources and skills for managing the software project.
- Development Environment: deals with how project development tasks can be achieved. Typical concerns are: operating system, software development tools, project time, and how tasks are allocated in developing the project.

4 RESULTS AND DISCUSSION

4.1 Study selection results

In total, after the application of the search string to the digital libraries, we identified 268 papers. The list of identified studies in the initial selection phase was comprised of 45 articles. After the full text reading during the final selection phase 12 studies were finally selected. Table 2 summarizes the results per each of the digital libraries used.

Table 2: Search Results

Digital Library	Returned studies	Initial selection	Final selection
IEEE Xplore	54	19	6
ACM	32	10	2
Science Direct	179	15	4
SpringerLink	3	1	0
Total	268	45	12

To answer our RQ, the data was extracted and synthesized from the 12 finally selected studies. The final list of cost attributes identified in this study comprises 39 cost attributes related to GSD project. The attributes have been classified into the three categories and are presented in the following subsections:

4.2 Cost attributes within “Product Engineering"

The software cost attributes derived from Product Engineering mainly relate to the complexity, reusability and reliability required to carry out the specified product, they are presented in Table 3

Nine software cost attributes have been identified to fit into this category, they deal with processes for handling product, description and documentation. The most commonly mentioned software cost attributes in this category is “Reusability". Reusing experiences and lessons learned in distributed software projects help to understand the characteristics and problems of distributed development, which might increase the project planning task.

Table 3: Software cost attributes related to “Product Engineering"

Cost attributes	Reference
Reusability	[3, 7, 11, 19, 20, 30, 31, 34]
Complexity of the software	[3, 11, 19, 20, 30]
Portability	[3, 11, 20]
Documentation	[19, 20, 30, 34]
Code size	[11, 19, 31]
Reliability	[19, 20, 22]
Requirement legibility	[3, 7]
Defect density	[7, 31]
maintainability	[20]

4.3 Cost attributes within “Project Personnel"

Personnel factors focus on the cultural fit, the closeness of team members and also the experience of the teams. It concerns educational level, language skills and knowledge management aspects of GSD projects. Nine software cost attributes that fall within this category have been identified in literature. They are presented in Table 4.

Software cost estimation in GSD projects includes specific cost overheads drivers and site-specific characteristics [14, 17]. The most reported cost drivers that fall into this category are associated to language skills and cultural differences which strongly affect distributed software development. Keil et al. [19] identified formal personnel factors, called “soft skills", to be a key success factor in distributed projects and provided mechanisms to coordinate cross-site work.

Table 4: Software cost attributes related to “Project Personnel"

Cost attributes	Reference
Language	[7, 11, 20, 25–27, 30, 31]
Cultural differences	[7, 11, 20, 25–27, 30, 31]
Communication	[7, 11, 20, 25–27]
Trust	[7, 11, 25–27]
Team structure	[7, 25–27]
Team size	[7, 11, 22, 31]
Personnel experience	[19, 20, 26, 31]
Development productivity	[7, 11, 19, 31]
personnel relationships	[19, 20, 27]

Another strongly cited cost attributes in this category is related to communication among teams working. Communication costs arise when there are communication related issues between the sender and recipient sites. Principally these are related to difficulties arising due to telephonic and electronic communication, instead of face-to-face communication [24].

4.4 Cost attributes within "Development Environment"

Twelve cost attributes have been identified in the "Development Environment" category. It is by far the category holding the largest number of cost attributes. They are presented in Table 5.

Table 5: Software cost attributes related to "Development Environment"

Cost attributes	Reference
Design and Technology Newness	[7, 11, 19, 26, 27, 31, 34]
Time zone	[7, 11, 20, 25–27, 30, 31]
Process model	[7, 11, 31]
Response delay	[7, 25]
Client involvement	[7, 11, 31]
Project effort	[7, 11, 22, 26, 31, 34]
Project management effort	[7, 11, 19, 26, 31, 34]
Task allocation	[11, 24, 26]
Work pressure	[11, 20, 26]
Work dispersion	[11, 26, 31]
Requirement volatility	[19, 20, 27]
Travel cost	[34]

The most frequently mentioned cost attribute is "Design and Technology Newness". Lack of knowledge on the required new technologies and lack of experience on the used effort estimation technique lead to wrong assumptions, which affect the accuracy of the cost estimates [16]. Another relevant cost attribute presented by the study was the "Time zone" that in many cases makes harder the process of hand off between sites, which also leads to wrong assumptions, compromizing the accuracy of the effort estimates.

5 THREATS TO VALIDITY

5.1 Construct Validity

Since the results of this study are drawn from an SLR, the search string used and the libraries queried can be identified as a threat to the construct validity. To minimize these threats we used the "PICO" method to extract initial search terms from the RQ and added synonyms drawn from similar research. We limited our SLR study to four research publication sources that are most common in software engineering (i.e., IEEE Xplore, ScienceDirect, ACM and SpringerLink). However, there are other related research sources that we did not consider in our study that may include relevant studies.

5.2 Internal Validity

Internal validity deals with data extraction and analysis. One author performed the data extraction and classification of the primary studies, and the other one reviewed the final results. The decision as to which data to collect and how to classify the papers was based on the judgment of the two authors conducting the review.

When conducting the SLR, we excluded no articles on the basis of their quality. Although some researchers might prefer to exclude articles of poor quality, including them served to clarify and develop the results of our SLR and allowed us to enrich our finding.

5.3 Conclusion Validity

Bias in the selection of software cost attributes in GSD context and data analysis may therefore affect the interpretation of the results. To mitigate this threat, every step in the selection, extraction and analysis of data was validated by means of the systematic process and periodic reviews carried out by the researchers involved in this work.

5.4 External Validity

External validity is related to the ability to generalize the findings of this study. It is unknown if the classification scheme used reflects the actual situation in industry. Interviews or surveys of industry practitioners involved in GSD projects to validate results were not conducted and represents one of the limitations of this study.

6 CONCLUSIONS

GSD is establishing itself as a mainstream paradigm of producing software, despite initial challenges faced by software companies. Project managers are quickly drawn to the benefits offered by such practice but lack the corresponding capabilities to handle relevant challenges. software companies especially need to be aware of these threats to their project success. In fact, Researchers and practitioners must continuously learn from failed past projects and aim to establish practices that produce better outcomes for the project.

In this paper, we proposed an up to date framework encompassing 30 cost attributes inherent to GSD projects derived from an SLR comprising 12 primary studies. The framework is divided into three categories based on the SEI taxonomy. The framework is intended for project managers engaging in the development of distributed software projects. However, this framework can only serve as a base for software cost area and must be adapted for each project. For future work, the authors intend to assess the relevancy of the items presented in the framework using a survey of researchers and industry practitioners.

ACKNOWLEDGMENTS

This work was conducted within the research project MPHR PPR1/09-2015-2018 . The authors would like to thank the Moroccan MESRSFC and CNRST for their support.

REFERENCES

[1] Aini Aman and Brian Nicholson. 2003. The process of offshore software development. In *Organizational Information Systems in the Context of Globalization*. Springer, 201–216.
[2] Lefteris Angelis, Ioannis Stamelos, and Maurizio Morisio. 2001. Building a software cost estimation model based on categorical data. In *Proceedings Seventh International Software Metrics Symposium*. IEEE, 4–15.
[3] Mohammad Azzeh. 2013. Software cost estimation based on use case points for global software development. In *2013 5th International Conference on Computer Science and Information Technology*. IEEE, 214–218.
[4] Barry Boehm, Bradford Clark, Ellis Horowitz, Chris Westland, Ray Madachy, and Richard Selby. 1995. Cost models for future software life cycle processes: COCOMO 2.0. *Annals of software engineering* 1, 1 (1995), 57–94.
[5] Barry W Boehm, Maria H Penedo, E Don Stuckle, Ronald D Williams, and Arthur B Pyster. 1984. A software development environment for improving productivity. *Computer* 6 (1984), 30–44.

[6] Pearl Brereton, Barbara A Kitchenham, David Budgen, Mark Turner, and Mohamed Khalil. 2007. Lessons from applying the systematic literature review process within the software engineering domain. *Journal of systems and software* 80, 4 (2007), 571–583.

[7] Ricardo Britto, Vitor Freitas, Emilia Mendes, and Muhammad Usman. 2014. Effort estimation in global software development: A systematic literature review. In *2014 IEEE 9th International Conference on Global Software Engineering*. IEEE, 135–144.

[8] Marvin J Carr, Suresh L Konda, Ira Monarch, F Carol Ulrich, and Clay F Walker. 1993. *Taxonomy-based risk identification.* Technical Report. CARNEGIE-MELLON UNIV PITTSBURGH PA SOFTWARE ENGINEERING INST.

[9] Saad Yasser Chadli, Ali Idri, Joaquín Nicolás Ros, José Luis Fernández-Alemán, Juan M Carrillo Gea, and Ambrosio Toval. 2016. Software project management tools in global software development: a systematic mapping study. *SpringerPlus* 5, 1 (2016), 2006.

[10] Manal El Bajta. 2015. Analogy-based software development effort estimation in global software development. In *Global Software Engineering Workshops (ICGSEW), 2015 IEEE 10th International Conference on*. IEEE, 51–54.

[11] Manal El Bajta, Ali Idri, José Luis Fernández-Alemán, Joaquin Nicolas Ros, and Ambrosio Toval. 2015. Software cost estimation for global software development a systematic map and review study. In *Evaluation of Novel Approaches to Software Engineering (ENASE), 2015 International Conference on*. IEEE, 197–206.

[12] Manal El Bajta, Ali Idri, Joaquín Nicolás Ros, José Luis Fernández-Alemán, Juan Manuel Carrillo de Gea, Félix García, and Ambrosio Toval. 2018. Software project management approaches for global software development: a systematic mapping study. *Tsinghua Science and Technology* 23, 6 (2018), 690–714.

[13] Manal El Bajta, Ali Idri, Joaquín Nicolas Ros, José Luis Fernandez-Aleman, and Ambrosio Toval. 2017. Estimation of Costs and Time for the Development of Distributed Software. In *Software Project Management for Distributed Computing*. Springer, 25–42.

[14] Manal El Bajta, Ali Idri, Joaquín Nicolás Ros, José Luis Fernández-Alemán, and Ambrosio Toval. 2017. Software Cost Attributes in Global Software Development Projects: A Survey. In *Proceedings of the 9th International Conference on Information Management and Engineering*. 96–101.

[15] Fred J Heemstra. 1992. Software cost estimation. *Information and software technology* 34, 10 (1992), 627–639.

[16] Mohamed Hosni and Ali Idri. 2017. Software effort estimation using classical analogy ensembles based on random subspace. In *Proceedings of the Symposium on Applied Computing*. 1251–1258.

[17] Ali Idri, Mohamed Hosni, and Alain Abran. 2016. Improved estimation of software development effort using Classical and Fuzzy Analogy ensembles. *Applied Soft Computing* 49 (2016), 990–1019.

[18] Staffs Keele et al. 2007. *Guidelines for performing systematic literature reviews in software engineering.* Technical Report. Technical report, Ver. 2.3 EBSE Technical Report. EBSE.

[19] Patrick Keil, Daniel J Paulish, and Raghvinder S Sangwan. 2006. Cost estimation for global software development. In *Proceedings of the 2006 international workshop on Economics driven software engineering research*. ACM, 7–10.

[20] Ansgar Lamersdorf, Jurgen Munch, Alicia Fernandez-del Viso Torre, Carlos Rebate Sánchez, and Dieter Rombach. 2010. Estimating the effort overhead in global software development. In *Global Software Engineering (ICGSE), 2010 5th IEEE International Conference on*. IEEE, 267–276.

[21] Raymond J Madachy. 1995. Knowledge-based risk assessment and cost estimation. *Automated Software Engineering* 2, 3 (1995), 219–230.

[22] Sajjad Mahmood, Sajid Anwer, Mahmood Niazi, Mohammad Alshayeb, and Ita Richardson. 2017. Key factors that influence task allocation in global software development. *Information and Software Technology* 91 (2017), 102–122.

[23] Tridas Mukhopadhyay, Steven S Vicinanza, and Michael J Prietula. 1992. Examining the feasibility of a case-based reasoning model for software effort estimation. *MIS quarterly* (1992), 155–171.

[24] Nanjangud C Narendra, Karthikeyan Ponnalagu, Nianjun Zhou, and Wesley M Gifford. 2012. Towards a formal model for optimal task-site allocation and effort estimation in global software development. In *2012 Annual SRII Global Conference*. IEEE, 470–477.

[25] Mahmood Niazi, Sajjad Mahmood, Mohammad Alshayeb, Abdul Majid Qureshi, Kanaan Faisal, and Narciso Cerpa. 2016. Toward successful project management in global software development. *International Journal of Project Management* 34, 8 (2016), 1553–1567.

[26] Mahmood Niazi, Sajjad Mahmood, Mohammad Alshayeb, Mohammed Rehan Riaz, Kanaan Faisal, Narciso Cerpa, Siffat Ullah Khan, and Ita Richardson. 2016. Challenges of project management in global software development: A client-vendor analysis. *Information and Software Technology* 80 (2016), 1–19.

[27] Srinivas Nidhra, Muralidhar Yanamadala, Wasif Afzal, and Richard Torkar. 2013. Knowledge transfer challenges and mitigation strategies in global software development—A systematic literature review and industrial validation. *International journal of information management* 33, 2 (2013), 333–355.

[28] Gerard O'Regan. 2015. Software engineering institute (sei). In *Pillars of Computing*. Springer, 195–205.

[29] Linda Hutz Pesante. 2003. Software engineering institute (SEI). (2003).

[30] S Ramacharan and K Venu Gopala Rao. 2016. Software Effort Estimation of GSD Projects Using Calibrated Parametric Estimation Models. In *Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies*. 1–8.

[31] Narayan Ramasubbu and Rajesh Krishna Balan. 2012. Overcoming the challenges in cost estimation for distributed software projects. In *2012 34th International Conference on Software Engineering (ICSE)*. IEEE, 91–101.

[32] Darja Šmite. 2005. A case study: coordination practices in global software development. In *International Conference on Product Focused Software Process Improvement*. Springer, 234–244.

[33] Darja Šmite and Juris Borzovs. 2006. A framework for overcoming supplier related threats in global projects. In *European Conference on Software Process Improvement*. Springer, 50–61.

[34] Darja Šmite, Ricardo Britto, and Rini Van Solingen. 2017. Calculating the extra costs and the bottom-line hourly cost of offshoring. In *2017 IEEE 12th international conference on Global Software Engineering (ICGSE)*. IEEE, 96–105.

[35] Anselm Straus and Juliet Corbin. 1990. Basics of qualitative research: Grounded theory procedures and techniques.