Dynamics of Collaborative Work in Global Software Development Environment

Fatma Cemile SERÇE[1], Ferda Nur ALPASLAN[2], Kathleen SWIGGER[3], Robert BRAZILE[4], George DAFOULAS[5], Victor LOPEZ[6]

[1]*Department of Information Systems Engineering, Atilim University, 06836, Ankara, Turkey*
Tel: +90 312 586 8375, Fax: + 90 31 2586 8091, Email: fcserce@atilim.edu.tr

[2] *Department of Computer Engineering, Middle East Technical University, 06531, Ankara, Turkey*
Tel: +90 312 210 5596, Fax: + 90 312 210 5544, Email: alpaslan@ceng.metu.edu.tr

[3] *Department of Computer Sciences and Engineering, University of North Texas, 311366, Denton, USA*
Tel: +1 940-565-2817, Fax: +1 940-565-279, Email: kathy@cs.unt.edu

[4] *Department of Computer Sciences and Engineering, University of North Texas, 311366, Denton, USA*
Tel: +1 940 565 4176, Fax: +1 940-565-279, Email: brazile@cs.unt.edu

[5] *School of Computing Science, Middlesex University, London, UK*
Tel: + 44 208 411 4402, Fax: +44 208 411 2332, Email: g.dafoulas@mdx.ac.uk

6 Facultad de Ingeniería de Sistemas Computacionales, Universidad Tecnologica de Panama, 0819-07289, Panama
Tel: +507 560 3603, Fax: +507 560-3624, Email: victor.lopez@utp.ac.pa

Abstract: This study aims to explore the dynamics of collaborative work in global software development projects. The study explored the nature of collaboration, the patterns of collaborative behaviors in different tasks in computer science, and the impact of the tasks to the collaboration among students. Four different collaborative software development tasks were assigned to the globally distributes teams. The study used data from 230 students from five universities, namely Atilim University (Turkey), Middle East Technical University (Turkey), Universidad Tecnológica de Panamá (Panama), University of North Texas (US), and Middlesex University (UK). The findings involve the recommendations for building effective collaborative working environments and guidelines for building collaborative virtual communities.

1. Introduction

Software industry is increasingly moving towards global software development (GSD) model because of its offer of large multi-skilled labor forces in lower cost and ability to quickly transfer product between development sites. Because of the lower development costs, global software development seems to be a phenomenon of increasing importance. However, it is not a magic to promise to success. These cross-site and cross-cultural projects do not just happen. There are many issues that require careful examination and consideration [1, 2, 3].

Software development is a human, social, and organizational activity as well as technical activity [1, 5, 6]. Global software development is highly based on human communication, coordination and control. Being people as the important factor is emphasized by Tom

DeMarco and Timotyh Lister in their book named as “Peopleware”[7]. While software and hardware are made of replaceable modular components, software development teams are not [1]. This needs to address peopleware issues becomes more critical in GSD. We need to understand how human behaves and collaborates in online global software development environment and what factors affects this behavior.

Software development involves collaborative work. There are three major processes in software development: communication, coordination, and control [8]. Global software development goes through these processes by means of online collaborative working environments [9,10]. Globally distributed work groups use this online environment to build their team, keep in touch, and work on projects together. To achieve a satisfactory result from global software development work, it is necessary to have a good communication and collaboration among the team members. Collaborative working has many dynamics to be considered. Dillenbourg and Schneider[11] argue some conditions for effective collaborative work such as group composition, task features, and communication media.

This study focus on collaborative work practices in a specific global software development settings. We try to figure out the particular challenges and the ways people use for coping with them in this specific situation. This paper presents a one-year study working on factors that contribute to the collaboration among students enrolled at Atilim University (Turkey), Middle East Technical University (Turkey), Universidad Tecnológica de Panamá, University of North Texas, and Middlesex University (UK). 230 students from the five universities participated in four projects. The projects were done in two semesters in year 2008. The student groups were assigned different types of software development projects that required some type of collaboration.

The main purpose of the study is to explore the dynamics on collaborative work among the globally distributed software development teams. To achieve this, we have the following research questions:

- What is the nature of collaboration among globally distributed teams?
- What does contribute to the effective collaborative work among the students?

3. Methodology

230 students from the five universities participated in four projects in 2008. Two of the projects were conducted with students in the United States and the United Kingdom. One of the project involved students from Turkey, Panama, and the US. The other project was conducted with the students from Turkey and US. In each project, there were randomly assigned teams, with approximately two or three students from each participating university. Since the universities are in different time-zones and having different academic calendar, the researchers first discussed the common date intervals for collaboration. Two of the projects were in spring semester and the other two were in fall semester of 2008. There were 43 globally distributed teams formed in the study. Each project setting has something in common, but also something in different.

3.1 Subjects, Tasks and Data Collection

The language for communication within the project teams was English, although the language for most of the co-located country teams in Panama and Turkey was their own native language.

Each project and their overall design are explained briefly in the following.

Project 1 : The major objective of the task was to teach students how to design and query a database. Thus, these collaborative projects focused on tasks such as creating a relational schema for a database, then writing queries to that same database. The task was

performed by MDX and UNT teams. There were 44 students from the two universities, and 10 teams having approximately 2 students from each university were formed. In this task, OAS!S+ a Virtual Learning Environment which is created through the customization of the WebCT Vista / Blackboard platform was used. This computer managed instructional software supports asynchronous communication tools such as forums, chats, and email. Totally 301 asynchronous messages were recorded. At the beginning of the task, a video-conferencing session was done with the participating students. The duration for the project was approximately 2 months in spring semester of 2008 (there were one week for a break).

Project 2: The student teams were assigned a code-intensive project, which consisted of developing a system that could be used to create and manage groups. Each country team was asked to deliver one of the components for the system and share that component with teammates in the other countries. The participating universities are University of North Texas (UNT), Universidad Tecnológica de Panamá (PTU), and Atilim University (AU). There were 86 students from the three universities. There were 10 teams formed, having 3 students from each university. The students from Turkey were all enrolled in a Java programming course at the Atilim University, while the students in the US were enrolled in a Human Computer Interface course at the University of North Texas. The students from Universidad Tecnológica de Panamá were not enrolled in a single course, but were recruited from several different project-oriented courses. There were 264 asynchronous messages recorded. At the beginning of the task, a synchronous chat event was done including all participants, and there were approximately 1800 chat messages recorded. For the communication environment, an open source platform learning management system called Online Learning and Training (OLAT). This tool also supports forums, chats, and emails as well as wikis. The duration for the task was almost one month in the spring semester of 2008.

Project 3: The task was to design, create and test a database software system for a car rental system case. This database development project includes tasks for determining the functionality of the system, designing and implementing a test database, writing SQL queries and implementing java program for testing, and producing a final report on the project. Each of these tasks describes a part of a deliverable. The dates and the tasks were pre-planned by the instructors. The participating universities were UNT, PTU and AU. There were 76 students from the three universities. There were 15 teams formed from the students from each university. OLAT was used for communication and collaboration environment for the teams. There were 260 messages recorded. The duration for the task was almost one month in fall semester of 2008.

Project 4: The main objective of the project task was to let the teams solve some of the problems in the field of artificial intelligence. There were three problems given and each team were required to solve it together. The project was between UNT and Middle East Technical University (METU). There were 34 students in eight teams. OLAT was used for communication and collaboration environment in the project. There were 166 messages recorded. The duration for the project was almost one month in fall semester of 2008.

3.2 Data Analysis

In 2001, Curtis and Lawson [12] reported on their exploratory study, which analyzed the online discussions of students who were engaged in a collaborative task. The primary data source for this study was the log of interactions that occurred while using an online discussion management system as well as the email messages that students sent to one another. The content of the students' messages and discussion were then analyzed for utterances indicative of collaborative behaviors, and a coding system was derived to help classify the different data elements. Since the goal of our current research project is to

determine ways to improve collaborative behaviors, we felt that the Curtis and Lawson coding scheme was the most appropriate coding scheme for our data (Table 1). The Curtis and Lawson coding system was based on a collaborative learning theory and was designed to characterize behaviors associated with positive social interdependence, as opposed to those behaviors linked to a more individualistic and competitive learning environment [12, 13]. By using this particular coding system, we should be able to determine the extent of collaborative behaviors that currently exist among our student groups, as well as measure the amounts of improvements that we may detect in future analyses, as we begin to experiment with different teaching strategies designed to improve collaboration. The actual analysis of the student transcripts involved coding the students‟ transcripts according to the Curtis and Lawson taxonomy [12]. Codes were assigned to utterances that indicated collaboration. Duplicate codes were assigned whenever an utterance indicated multiple collaborative behaviors.

Table 1 Coding Scheme and Behavior Categories [12]

Behavior Categories	**Behaviors**
Planning	Group Skills
	Organizing work
	Initiating Activities
Contributing	Help Giving
	Feedback Giving
	Exchanging Resources and Information
	Sharing Knowledge
	Challenging others
	Explaining or elaborating
Seeking Input	Help Seeking
	Feedback Seeking
	Advocating Effort
Reflection/Monitoring	Monitoring Group Effort
	Reflecting on medium
Social Interaction	Social Interaction, SI

In the study, cluster analysis is performed to investigate collaborative tasks. The primary purpose of cluster analysis is to identify the tasks based on their support on collaborative behaviors and to find patterns or characteristics among similar projects. The groups or clusters that result from this classification process should identify characteristics that maximally discriminate among the cases in different segments. The objective for this particular study was to identify distinct groups of collaborative tasks with similar collaborative behaviors, we can explore how these distinct patterns of tasks have group dynamics.

The clustering variables were each group's number of interactions devoted to the five interaction behaviors. Based on a review of clustering techniques, we chose a hybrid clustering method to identify the different groups. The hybrid clustering technique uses two methods namely k-means and Ward's hierarchical agglomerative clustering. The centers (or centroids) of each cluster are obtained first using Ward's method [14], a hierarchical cluster analysis technique that is said to be the most likely method to discover any underlying

cluster structure. The resulting centroids are then used as the initial seed points for the non-hierarchical k-means cluster analysis.

4. Findings

There were 43 globally dispersed teams including 240 students in total. On the average, there were 991 asynchronous messages recorded during the life time of the projects.

In Project 1, at the beginning of the semester a video-conferencing session was conducted. In Project 2, there was a synchronous event, chat session, at the beginning of the project. In this chat session, there were 1800 messages and almost 800 of them were codes as messages in "social interaction" behavior.

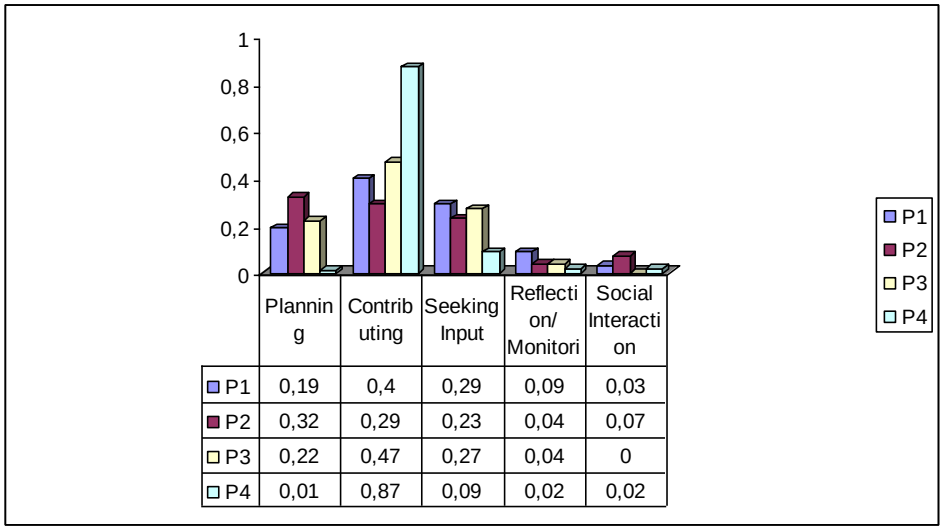
In order to investigate the nature of collaboration interaction in global software learning teams, the asynchronous messages recoded by learning management systems was content-analyzed upon completion of the projects. To conduct the content analysis, each chat/forum discussion messages was coded to determine the overall number of the communication behaviors such as planning, contributing, seeking input, reflection and monitoring, and social interaction. Individual codes were assigned to postings that indicated specific types of behavior. If the behavior was not present in a communication incident, it was assigned a score of 0; conversely, if a communication behavior(s) was the purpose of a posting or chat, the posting was assigned the code or codes for that behavior. As a reliability check, a second coder analyzed the same discussions and chats. Inter-rater reliability between coders for the interactions behaviors was acceptable (.84)

After content analysis, the utterances of each project in online collaboration are figured out as shown in Figure 1. As it can be seen, social interaction behavior category is the one that has the least percentage for almost all projects except for Project 2. This might be because that Project had chat session as a synchronous activity and this might help project members to keep socializing during their asynchronous communication.

In the study, cluster analysis was used to identify tasks based on their support on collaborative behaviors and to find patterns or characteristics among similar projects. According to the cluster analysis, the agglomeration schedule generated from Ward's method suggests a four-cluster solution. Figure 2 shows the clusters and project's category of cluster.

Cluster 1 includes planning, contributing and seeking input behaviors in almost equal proportion. There are almost no record of reflection and monitoring in Cluster 1 and Cluster 4. The global collaborative workers in Cluster 1 and Cluster 4 spent much more time on planning. Cluster 3 has a very different pattern of collaboration. Contributing behavior is dominant in cluster 3. It seems that the member in each project asked for assistance or feedback from others. The students in Cluster 3 spent time on monitoring group effort and give comments about the effectiveness of the medium. Cluster 1, 2 and 3 spent the least amount of time on social interaction. The global software learners in Cluster 4 spent much more time on socializing.

Having identified the behavioural patterns, we, then, examined the relationship between collaboration and project settings (Figure 3). It seems that Project 2 has somewhat a proportion in all clusters. It means that it encourages some of the project teams to have almost all collaborative behaviors such as planning, contributing, seeking input, reflection and monitoring and social interaction. Since there were synchronous activities in both Project 1 and Project 2, only these two projects in Cluster 4, which has property of having the highest record of social interaction among the other clusters. Project 4 seems to be the project promoting the most contributing behavior among the other projects, because it appears mainly in Cluster 3, which has the highest contributing behavior.



	Planning	Contributing	Seeking Input	Reflection/ Monitoring	Social Interaction
P1	0,19	0,4	0,29	0,09	0,03
P2	0,32	0,29	0,23	0,04	0,07
P3	0,22	0,47	0,27	0,04	0
P4	0,01	0,87	0,09	0,02	0,02

Figure 1 Proportions of Collaborative Behaviors in each Project

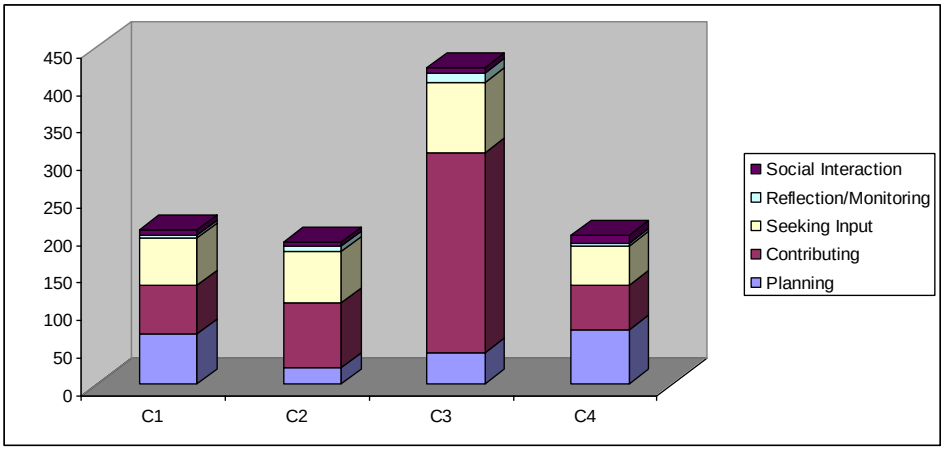


Figure 2 Proportions of Collaborative Behavior in each Cluster

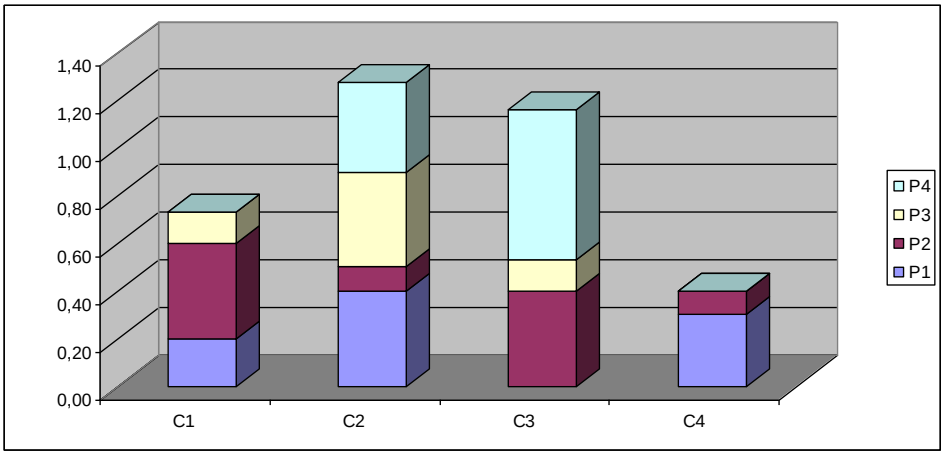


Figure 3Projects in Clusters

5. Discussions and Conclusions

230 students from the five universities participated in four projects in 2008. Project 1 includes designing and querying a database. 44 students from two universities were asked to create a relational schema and then write queries to that same database. At the beginning of the project, a video conferencing session was held. From the findings of the study, it can be understood that this video-conferencing event help socializing among students.

A total of 86 students participated in Project 2. Students were randomly assigned to one of 10 groups. Each group contained approximately three co-located members and six dispersed members. This particular project required teams to develop a group management application. Similar to Project 1, at the beginning of the project, a synchronous chat event was done including all participants, and there were approximately 1800 chat messages recorded. This project includes both cooperative and collaborative working experience. Because each country team was asked to deliver one of the components for the system. Each local team were asked to develop individual components, a means of cooperation. After having all parts, they are asked to integrate all parts, which promoted collaboration. The findings of the study indicated that both cooperative and collaborative tasks together could have encouraged global teams to experience in each collaborative behavior category such as planning, contributing, seeking input, reflection and monitoring and social interaction.

The third project was to design and develop an online system for renting automobiles. A total of 70 students participated in this project. This database development project includes tasks for determining the functionality of the system, designing and implementing a test database, writing SQL queries and implementing java program for testing, and producing a final report on the project. The project deliverables were previously determined by the instructors. In this project, each country side has different roles such as project manager, developer, database designer and tester. The findings showed that this project also promoted the global teams to have different collaborative behaviors except social interaction. This might be because there were no synchronous events during the project.

Project 4 was about to ask global teams to solve some of the problems in the field of artificial intelligence. There were three problems given and each team were required to solve it together. There were 34 students in eight teams. This project was different from the other projects. In this project, the students were not supposed to design and develop a database and/or an application. They were asked to collaboratively solve some of the artificial intelligence problems. So, they were need to collaborate in problem solving. However, in reality this was not that case. Almost all members sent individual solutions and the the instructor sent feedback about its correctness. Although this project seems to have the highest proportion of contributing behavior, almost all contributing behaviors were in the form of feedback giving. There were almost no sharing knowledge, exchanging resources behaviors in the contributing behavior category. This means that this proejct did not promote collaboration effectively.

Software development teams pass through certain software development phases such as project plan, analysis, design, implementation, test, and integration. Project planning requires determining schedules, roles and responsibilities at least. That means there should be records of initiating activities and organizing behaviors under planning behavior categories. For the other phases, the team members need to come together, discuss about the requirement, exchange ideas and resources, give feedbacks about the other design, giving help about using any one of design tool, help seeking to run the application for test etc. For the project management issues, it is necessary to monitor the group's effort. The project manager should give feedback to team, should follow if they are on the schedule. In short, it is necessary to have all collaborative behaviors given in the coding scheme:

planning, contributing, seeking input, reflection and monitoring in a global software development project. There is one more behavior social interaction. It is crucial in global setting[15]. When there are people in the virtual working environment, then it is necessary to have social activities. Otherwise, it will be an artificial virtual team, having members not knowing each other, not understanding each other, not considering the other all.

6. Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 0705638. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. We also wish to thank the students who participated in the study, and the many colleagues (in all four countries who helped make this research possible.

References

[1] Avram, G. "Of Deadlocks and Peopleware - Collaborative Work Practices in Global Software Development", ICGSE '07: *Proceedings of the International Conference on Global Software Engineering*, 91-102, IEEE Computer Society, Washington, DC, USA, 2007

[2] Ågerfalk, P.J. , Fitzgerald, B., Olsson, H., and Conchúir, E. "Making Globally Distributed Software Development a Success Story, Benefits of Global Software Development: The Known and Unknown", *Lecture Notes in Computer Science*, Volume 5007, Springer Berlin / Heidelberg, 2008

[3] Lings,B., Lundell,B. and Agerfalk,P.J. and Fitzgerald, B. "A reference model for successful Distributed Development of Software Systems", ICGSE '07: *Proceedings of the International Conference on Global Software Engineering*, pp. 130-139, IEEE Computer Society, 2007.

[4] Massey, A.P., Montoya-Weiss, M.M. and Hung,Y. "Because Time Matters: Temporal Coordination in Global Virtual Project Teams". *Journal of Management Information Systems*, 19, 4, pp. 129-155, M. E. Sharpe, Inc, 2003.

[5] Herbsleb, J.D., and Moitra, D. "Global Software Development", *IEEE Software*, IEEE, March/April 2001.

[6] Kommeren,R., Parviainen,P. "Philips experiences in global distributed software development", *Empirical Software Engineering*, 12, 6, 647-660, Kluwer Academic Publishers, Hingham, MA, USA, 2007.

[7] DeMarco, T. and Lister, T. "Peopleware: productive projects and teams", Dorset House Publishing, New York, 1987.

[8] Holmstrom,H., Conchuir,E.O., Agerfalk, P.J., Fitzgerald,B. "Global Software Development Challenges: A Case Study on Temporal, Geographical and Socio-Cultural Distance", ICGSE '06: *Proceedings of the IEEE international conference on Global Software Engineering*, 3-11, IEEE Computer Society, 2006.

[9] Lanubile, F., Damian, D., Oppenheimer, H.L., "Global software development: technical, organizational, and social challenges", *SIGSOFT Softw. Eng. Notes*, 28, 6, ACM, 2003.

[10] Gotel, O. Scharff, C. and Seng, S. "Preparing Computer Science Students for Global Software Development". In *Proceedings of the IEEE Annual Frontiers in Education Conference, Borders: International, Social & Cultural*, San Dieago, Californai, USA, 2006.

[11] Dillenbourg, P., Baker, M., Blaye, A. & O'malley, C. "The Evolution Of Research On Collaborative Learning", In E. Spada & P. Reiman (Eds) Learning in Humans and Machine: Towards an interdisciplinary learning science. Oxford: Elsevier, pp. 189-211, 1996.

[12] D. Curtis and M. Lawson, "Exploring Collaborative On-line Learning", *JALN*, 5, 1, 2001.

[13] R. Marra, J. Moore, and A. Klimczek, "A Comparative Analysis of Content Analysis Protocols for Online Discussion Forums", *Educational Technology Research and Development*, 52, 2, 2004, pp. 23 - 40.

[14] Afifi, M., V. Clark, and S. May, "Computer-Aided Multivariate Analysis" (4th Edition) Chapman & Hall/CRC Press, 2004.

[15] Kotlarski, J. and Oshri, I., "Social ties, knowledge sharing and successful collaboration in globally distributed system development projects" , *European Journal of Information Systems*, 14, pp 37-48, 2005.