Understanding collaboration in Global Software Engineering (GSE) teams with the use of sensors

Introducing a multi-sensor setting for observing social and human aspects in project management

Georgios Dafoulas
Cristiano Maia
Almaas Ali
Juan Augusto
Computer Science Department
Middlesex University
London, United Kingdom
g.dafoulas@mdx.ac.uk
c.maia@mdx.ac.uk
a.ali@mdx.ac.uk
j.augusto@mdx.ac.uk

Victor Lopez-Cabrera
Computer Systems Engineering Department
Technological University of Panama
Panama
victor.lopez@utp.ac.pa

Abstract—This paper discusses on-going research in the ways Global Software Engineering (GSE) teams collaborate for a range of software development tasks. The paper focuses on providing the means for observing and understanding GSE team member collaboration including team coordination and member communication. Initially the paper provides the background on social and human issues relating to GSE collaboration. Next the paper describes a pilot study involving a simulation of virtual GSE teams working together with the use of asynchronous and synchronous communication over a virtual learning environment. The study considered the use of multiple data collection techniques recordings of SCRUM meetings, design and implementation tasks. Next, the paper discusses the use of a multi-sensor for observing human and social aspects of project management in GSE teams. The scope of the study is to provide the means for gathering data regarding GSE team coordination for project managers including member emotions, participation pattern in team discussions and potentially stress levels.

Keywords—Global Software Engineering; social and human aspects; software engineering project management; collaboration; communication; sensor data in software development cooperation;

I. INTRODUCTION

This paper merges experiences from previous research in GSE and in particular conducting empirical studies in the way GSD teams collaborate and communicate with the use of sensor-generated data for managing such teams. There is a significant volume of literature in relation to findings from work in GSE, ranging from putting together the necessary infrastructure, setting up pilot studies, observe collaboration and communication patterns as well as investigation in human issues relating to virtual team formation and remote work.

The paper provides an excellent opportunity for investigating how sensor data can be used for the understanding how GSE teams collaborate, how their members communicate and how their work is being coordinated. The discussion begins with a literature review in a number of related fields before focusing on previous lessons learnt from conducting observations of student teams engaging in GSD projects. The main part of the paper provides a detailed review of how a pilot study on GSE student work was set up and emphasizes on the use of multi-sensor settings. The scope of the paper is to discuss issues relating to the use of sensor-generated data for managing GSE projects and better understand social and human aspects of GSD.

II. LITERATURE REVIEW – COLLABORATION IN GSE

Our work is based on years of experience with global software engineering student teams, and the deployment of pilot studies involving collaboration over distance, different time zones and cultural differences. Previous work involved the investigation of team communication and collaboration patterns during small-scale Global Software Development (GSD) project. Emphasis was given on the interactions between team members, the exchange of information between teams collaborating on specific tasks and the interactions taking place with the support of synchronous and asynchronous communication. Our work emphasized primarily on three areas of work, namely (i) cross-site collaboration in GSE, (ii) team building and teamwork in GSE scenarios and (iii) social and human aspects in GSD. This paper presents the latest research considerations involving the use of sensors for gathering data essential for project management decisions in GSE. Although the paper discusses work in progress on setting up multi-sensor settings for the coordination of GSE teams, the scope of this
research is to provide guidance for establishing practices for ‘smart’ GSD project management. This should be achieved by supporting decision making for GSD coordination and communication with data collected from sensors monitoring human behavior.

One of the key issues in GSE education is to provide learners with the opportunity to experience realistic scenarios so they can obtain the necessary skillset. There are several examples in the relevant literature on how GSE skills are practices in learning settings. Such skills may range significantly; therefore it is necessary to adopt a variety of methods. For example Paasivaara et al [13] in order to assess learning of GSE competencies they employed a mixed-method approach including “post-course interviews, pre-, post-course and iteration questionnaires, observations, recordings of daily scrums as well as collection of project asynchronous communication data”. It is also important to investigate additional factors affecting group performance, such as “gender, age, cultural diversity, previous work experience, and the degree to which work is equitably shared among team members as possible factors affecting success” [11]. There are several examples of work sharing experiences in setting up distributed software engineering courses and sharing good practices. It is important to note that teaching intentions vary across institutions and focus may shift across a number of areas such as project management, requirements engineering & quality assurance, architecture, and implementation [6].

With respect to team building and teamwork in GSE scenarios, a detailed literature review conducted in 2012 [14] identified 74% of relevant publications focusing on distributed team and project management. The vast majority of work in the area is concerned with the main challenges associated with GSD [8] such as (i) creating overlap in time between different sites, (ii) geographical distance challenge team spirit, i.e. ‘teamness’ and (iii) the inherent challenge of creating mutual understanding between people with different backgrounds. Furthermore, GSE has a significant impact on requirements understanding [10] due to its special nature that translates into cultural differences, loss of communication richness, loss of teamness, time zone differences, coordination breakdown and geographic dispersion.

Apart from issues relating to cross-site collaboration, collaboration and communication, it is important to consider social and human aspects in GSE. The added complexity of GSD projects relating to human issues comes from (i) lack of common understanding of goals and requirements assigned, (ii) difficulties in communication (members are geographically separated), (iii) bottlenecks and problems in project execution (Variety of processes, management mechanisms, and associated skills/competencies) and (iv) ineffective management of knowledge sharing [7]. According to Misra et al [12] GSD challenges relating to personnel include (i) communication, (ii) knowledge management, (iii) coordination, (iv) collaboration, (v) socio-cultural distance (lack of group awareness), and (vi) lack of trust.

Interestingly enough sensor-based project management provides an excellent opportunity for increasing the effectiveness of dealing with social aspects in GSE.

Ara et al [1] propose a sensor-based project management process, which uses “continuous sensing data of face-to-face communication, was developed for integration into current project management processes”.

### III. BACKGROUND WORK IN GSD

Our previous work involved several years of investigation in the field of GSD and especially in the way virtual teams are formed, and their patterns of collaboration and communication [5]. Previous work involved more than thirty pilots involving GSD teams from six institutions involving four countries (UK, USA, Panama and Turkey). Our studies emphasized on understanding the way students would interact during software development activities ranging from design and analysis tasks to coding and interface testing. Focus was also on assessing communication patterns during synchronous and asynchronous discussions using text coding. Amongst the research objectives was the investigation of how culture as well as time zone differences affected the way teams collaborated in short projects.

Furthermore previous work focused on analyzing the collaborative interaction patterns of global software development learning teams composed of students from different countries. In particular, qualitative and quantitative analysis methods were used to determine the differences between a group’s communication patterns in asynchronous versus synchronous communication mode. K-means clustering with the Ward method was used to investigate the patterns of behaviors in distributed teams [15].

Another key consideration of previous work was assessing e-learning, virtual team and GSD platform infrastructures available for supporting such work in a feasible and sustainable way [3]. Emphasis was on determining aspects of the necessary infrastructure to support pilot studies in GSE. The scope of the research was twofold. First, it was important to find a feasible method for replicating virtual team scenarios and GSD tasks for software engineering students. Second, it was important to provide a framework for GSE engineering allowing learners to interact in realistic GSD scenarios and experience social aspects of such work. The following two sections will briefly present two pilot studies based on the creation of such scenarios. Our work was also driven towards considering stress assessment at the workplace [18].

### IV. ENHANCING SUPPORT OF GSD TEAMS

Based on the experiences from investigating GSE team communication and collaboration patterns, we shifted our focus on two new areas of concern. The first aim was to enhance project management in GSE with the use of smart technology. The second aim was to use visual analytics for facilitating the evaluation of team performance with a number of key indicators.

As part of this new research direction, a pilot study was designed with the participation of students from seven
universities. Students from three Egyptian and three Palestinian institutions collaborated with students from a Panamanian university.

There were twelve sub-teams organised consisting of members from the same university. In total 78 students participated in the study, in three teams. The project lasted for two weeks and there were three milestones every week, each having specific deliverables corresponding to conceptual modeling and database design tasks. Participants communicated through an online synchronous chat and a threaded discussion supporting asynchronous interactions. Every morning the four sub-teams that were part of each main team had to get together and perform a SCRUM meeting. Twice a week a video conferencing session was taking place so team members would receive guidelines for the rest of the project.

The communication server logs were analyzed to produce a detailed dashboard for individual, sub-team and team activity throughout the project. Team managers obtained an overall performance assessment of their team, while individuals could obtain a progress assessment of their own efforts.

Selected members from two of the three were provided with four wristbands, which were used, from the team managers and three nominated members. The heart rate monitors were triggering a stress monitoring survey twice a day based on the team members reaching a pre-determined threshold. Members of the third team would receive the same stress monitoring survey once it was triggered by any one of the other two teams.

The information from the heart rate monitors (wristbands) was analyzed to produce stress patterns of the nominated members. Team managers obtained an overall view of team stress levels, while individuals could obtain an assessment of their own efforts.

Selected members from two of the three were provided with four wristbands, which were used, from the team managers and three nominated members. The heart rate monitors were triggering a stress monitoring survey twice a day based on the team members reaching a pre-determined threshold. Members of the third team would receive the same stress monitoring survey once it was triggered by any one of the other two teams.

The information from the heart rate monitors (wristbands) was analyzed to produce stress patterns of the nominated members. Team managers obtained an overall view of team stress levels, while individuals could obtain an assessment of their own efforts. Once a certain threshold was reached from the heart rate measurements, the stress monitoring survey app triggered a questionnaire that was sent to all team members. The responses from the stress monitoring survey were displayed through Tableau [4].

V. INVESTIGATING GSE PROJECT MANAGEMENT

The second pilot study focused on investigating team coordination and communication with the aid of a multi-sensor setting. The scenario was based on the observation of teams consisting of 4-6 members working on a strategic information systems course. Student teams had to reach consensus for the content of a brief presentation providing an executive summary of their project work and reflection on individuals’ contribution to teamwork.

The study involved final year students of Information Technology and Business Information Systems courses and took place in a dedicated ‘smart lab’. The laboratory is in the form of a smart-home including various rooms with sensors installed. The experiment used two rooms, the first room for a project management meeting and the second room for a group presentation.

The experiment consisted of six phases as follows:
- Induction – one week prior to the experiment the students were given a tour of the smart lab and were acquainted with the research study, while one hour before the monitoring students were provided with the research briefing documentation, consent forms and detailed explanation of what information was collected and the sensors used during the study.
- Reflection – participants were given a few minutes to consider their contribution to the teamwork over the past ten weeks and assess their ability to present certain topics.
- Coordination – each team appointed a team leader for the project meeting who was responsible to ensure that each member would provide sufficient input in the discussion and that the team would reach consensus.
- Communication – all members had to exchange their views on which topics each member should present, working towards team consensus.
- Presentation – the team had to present a selection of topics following the agreed plan during the project management meeting.
- Feedback – the team was provided with immediate feedback on their presentation, the process they
followed to reach consensus and indicative measurements from the different sensors.

The experiment attempts to provide several sources of additional data that can help GSE team leaders to make decisions in relation to team meeting effectiveness. The benefit for future GSE work could include the ability to identify actual participation for each member in team decisions, impact of team meetings on individuals and evidence for individual contributions on teamwork.

Participants were provided with a form when they entered the smart lab, and were initially requested to determine the number of topics their team should present during the 60-90 seconds available for the team presentation. Next, participants were required to select from a list of eighteen (18) topics which ones should be presented. Finally, participants should identify the team member who should present each of the selected topics. Once the individual forms were completed, the team coordinator started the team meeting, having 10 minutes to reach consensus. The coordinator had a second form, which was used to confirm the topics selected for the team presentation, identify the members who would present each topic and determine the order of the topics. Table I lists the concepts associated with GSE team coordination that are part of this research study.

<table>
<thead>
<tr>
<th>Concepts</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration</td>
<td>The time it took the team to reach consensus</td>
</tr>
<tr>
<td>Participation</td>
<td>Evidence of involvement in team communication.</td>
</tr>
<tr>
<td>Presentation</td>
<td>The proportion of each member’s presentation part.</td>
</tr>
<tr>
<td>Concern</td>
<td>Evidence of anxiety or stress during discussions.</td>
</tr>
<tr>
<td>Disagreement</td>
<td>Frequency of discussion debates and disagreement.</td>
</tr>
<tr>
<td>Emotion</td>
<td>Range of emotions during specific meeting milestones.</td>
</tr>
<tr>
<td>Contribution</td>
<td>Evidence of individual effort towards teamwork.</td>
</tr>
<tr>
<td>Misalignment</td>
<td>Difference between individual and team decisions.</td>
</tr>
</tbody>
</table>

The experiment involved 13 student groups with a total of 53 participants. The aim of the study was to investigate social and human aspects of software engineering tasks based on previous experience on setting up student GSE projects. The study also aimed to investigate the validity of the data collection approaches used with a number of sensors in order to understand the way GSE teams were coordinated and their members communicated. Our hypothesis is that sensor data can be used to understand team cohesion, team dynamics, individual contribution to team decisions, states of emotional arousal and possible associations between the concepts identified in table I.

VI. MULTI-SENSOR SETTINGS FOR MONITORING GSD

As mentioned earlier, the second study discussed in this paper was conducted in a smart-lab. The facility is a sensor-equipped home and was selected to provide teams with an alternative environment to their usual laboratory. The aim was for participants to have minimum disruption and feel the entire facility was available for their meeting and presentation. Participants were allowed to use either the living room or kitchen areas to reflect on their individual contributions and fill in their forms. They would then move to one of the study spaces for their project meeting before delivering their presentation in the living room area.

Several sensors and data collection techniques were considered in preparation for the pilot study. Although heart rate monitors were used in the first pilot study it was decided not to use them, as there was no sufficient time to monitor individual heart rate patterns in order to provide a comparable baseline. Galvanic Skin Response sensors were considered for monitoring perspiration as an alternative to heart rate monitors that were suitable for the team meeting setting, as there was minimum movement across the room. The use of a portable polygraph to record physiological measures such as respiration, pulse, blood pressure and skin conductivity, although very useful was considered as intrusive and impossible to use without affecting the project meeting process. Heat sensor cameras were also considered for the presentation recordings and it was decided for use in a follow up pilot. Another possible source of data was the use of eye tracking software for measuring blink rate and direction of gaze. It was decided that this sensor would be better used for a scenario that would require one-to-one communication where two members would face each other when negotiating. Furthermore, the fact that all members would be sitting around a meeting table meant that the use of a sociometer for measuring cooperation and collaboration in physical space would not be of use for the specific pilot. Researchers at MIT have worked on “using statistical pattern recognition techniques such as dynamic Bayesian network models we can automatically learn the underlying structure of the network and also analyze the dynamics of individual and group interactions” [2]. This work provides an excellent opportunity for understanding the structure of face-to-face interactions and possibly to replicate similar models for online interactions.

The sensors and recording devices used for the collection of data during the pilot study focused on the collection of video, audio, and physiological data. Emphasis was given on identifying the extent of individual contribution, while assessing individuals’ emotion. Future plans include further analysis of the sensor data with tagging of the meeting videos.

Audio data was collected using the Kinect for Xbox One, motion-sensing device. The device was programmed to collect the sound source angle (in degrees), and the direction that sound is arriving from a sound source. This enabled to observe the verbal participation of the individuals in the group. Audio data was collected during the team decision-making stage and presentation stage. This allowed determining individuals’ participation in the team meeting and the proportion of the presentation each member delivered. Our objective was to investigate whether team cohesion could be determined from team member contribution using audio behavior. This would be in line with existing work on analyzing group behavior within the context of cohesion and especially automatically estimating high and low levels of group cohesion [9].

Electrical conductance of the skin was measured using galvanic skin response sensors. Six such sensors were built, one for each team member. Each participant had to hold the sensor during the entire project meeting (on their thumb) measuring physiological and emotional arousal. This type of
sensor, although it has received criticism on the merit of its accuracy and the ability to provide an acceptable measure of stress, has been used in the past [16]. Our objective was to investigate possible association from different conductance of the skin at certain points of each meeting that could be perceived as stressful.

Further to the previous two sensors, each participant had a camera focusing on their facial expressions, connected to a laptop. Participants’ facial expressions were collected in real-time during the decision making process. Team members’ facial expressions were collected as input and returned set of emotions for each face as well as the bounding box for the face using Microsoft Face API. The possible emotions to be detected were anger, contempt, disgust, fear, happiness, neutral, sadness and surprise, which are universally communicated using facial expressions. During this preliminary investigation the team did not focus on the selection of different algorithms for the analysis of emotions, as the scope was to investigate the usefulness of this data stream rather than the accuracy of the emotion of each subject. Further work is planned to determine the level of accuracy for emotion detection based on specific algorithms. The objective was to associate expressions to particular emotions of team members during certain points during the decision making process and throughout the consensus meeting. In the past robust recognition of facial expressions from images and videos was still a challenging task due to the difficulty in accurately extracting emotional features [17]. However, “significant performance improvements due to the consideration of facial element and muscle movements” have improved the performance of facial expression recognition systems. Based on the preliminary data analysis we have conducted we are working towards using the same technique during video conferencing between remote team members when collaborating in scenarios similar to the one described in the first pilot study.

Using a camcorder video and audio of the group presentation was recorded and several pictures were taken. The objective was to tag videos for synchronizing the data collected from various sensors and also identify important milestones during the decision-making and presentation stages. Any notes, which participants had written on the given forms during the experiment were collected, to analyze input individuals had on the final decision. Figure 3 shows the setting of the decision-making project meeting and the set-up of the sensors, recording devices and laptops.

VII. USING SENSOR-DATA FOR PROJECT MANAGEMENT SUPPORT

Using sensor-data for supporting project management decisions requires analysis of the data set collected from each data sensor. Future work will focus on creating appropriate APIs connecting to data visualization applications like Tableau. The following figures demonstrate how sensor-data can lead to certain assumptions for aspects of the GSE team coordination and communication.

A. Investigating team members’ emotions during GSE collaboration

Figure 4 shows the pattern of facial expressions for group 8 consisting of five members. Several of the 13 groups demonstrated similar patterns where a significant number of the 53 participants had a neutral dominant expression. By observing the video files of the team meetings two key patterns emerged. First the students demonstrating neutral emotional patterns seemed to be the ones who participated the least in team discussions, denoting lack of interest. Second, there seemed to be a number of students who were less confident during negotiations and while being active listeners, preferred to adopt a more passive approach in the decision-making process, following the lead of dominant coordinators or other team members.

As shown in figure 5 there is an association between the emotions expressed by team members while trying to reach consensus. For example, at the time P5 expressed anger during negotiations, P2 and P3 demonstrated signs of sadness. There was also a dominant emotion of sadness throughout the meeting aligned to the difficulty in reaching an agreed presentation plan. The data collected from the emotion detection software can be used to support GSE project managers in a number of ways. One of the key issues in virtual teams and GSD projects is the lack of rich, face-to-face communication where mannerisms and non-verbal communication can help to interpret members’ views on decisions. Furthermore, collocated teams engaging in collaboration activities such as SCRUM meetings tend to focus on the message of each speaker rather than body language and facial expression. This input could be analyzed after the meeting to provide further information to the project manager about the state of each member. This additional information could be used to interpret certain patterns on individual performance, conflicts between members and even satisfaction levels.

Figure 3. Multi-sensor setting for GSE experiment photos.
B. Investigating individual contribution during GSE collaboration

During both the consensus meetings and the team presentations the use of Kinect helped to determine the source of audio. This type of sensor data was originally calculated in a quantitative way in order to determine those members who contributed the most during both tasks. However, future analysis will focus on making associations between certain video tags prompting changes to the person who contributes, as well as those timestamps where several members react on certain statements.

Figure 6 illustrates how four different members contributed to the consensus discussion. There is clear input from four input directions shown at 0, -400, -1000 and -2800. It is obvious that the person located across the Kinect device dominated the discussion.

Figure 7 shows how the application of the filter increases the confidence level (the lines become thicker) and we can see that the other three members of the first group provided very brief input. Each member provided a single input for a couple of seconds. This is not though the case for the second team that has several members contributing to the discussion. This team shows that there was more discussion during the decision-making process and that different views were taken under consideration.

This type of sensor data can provide a useful input for project managers who wish to determine whether there are certain dominant members in the team, to identify any lurkers or members demonstrating apathy and also evaluate whether decisions were reached after sufficient discussion and consideration.

Further analysis will help us to find associations between the individual levels of contribution in different tasks. For example, is the dominant person of the decision-making phase also dominant during the presentation? Future work will also attempt to find patterns between contributions in team meetings and actual input on team activities that is measured with a number of assessment criteria. For example is the dominant communicator also a hard working team member or a leading figure that frequently delegates work?
Interestingly enough only a few teams demonstrated a pattern where the appointed coordinator was the main contributor in both decision-making discussions and presentation.

At this stage it is important to note that scholars who wish to replicate the experiment presented in this study should assess whether the Kinect’s ‘confidence’ feature is reliable. It is important to check the video recordings of such meetings in order to ensure that when the ‘confidence’ value is 0 only ambient sound is recorded. If the set up is not appropriate it may be that members who are at a blind spot are not recorded, and the same may happen for those individuals who have a soft, calmer voice.

C. Use of GSR sensors in GSE collaboration

Finally, the use of GSR sensors provided input for the electrical conductance of the skin. For the majority of the 53 participants, the nature of the activity meant that there was no useful input. This was in line with our assumption that this type of sensor would be useful only in combination with other sources of input. However, certain teams and their members provided useful patterns during more in-depth discussions. As shown in figure 20, although three of the team members did not provide any indication of change during the communication, members P5 and P6 show similar patterns that are at certain points aligned to the pattern demonstrated by P2.

VIII. CONCLUSIONS AND FUTURE WORK

This paper discussed how to use sensors in order to collect data that can facilitate the role of GSE project managers. The discussion was primarily based on the lessons learnt from previous work, earlier use of sensors to record stress levels and stress awareness in GSD teams as well as ongoing research in social aspects of GSE. The paper shared experiences in setting up a multi-sensor observation space for GSE work and explained how the setting could be used to record data during decision-making and presentation tasks.

The information gathered from multiple sensors has helped the authors to create a data set that can be further analysed in order to facilitate decision making when managing such teams in the future. The use of sensors has helped in determining whether certain patterns can emerge from observing certain behaviours and reactions of individuals when participating in GSE projects. The authors are currently developing the dashboard to present the various manipulations and canonical transformations on the data to provide project managers and team leaders of GSE teams with a useful toolkit. The scope of this work is to ensure project managers have richer information when managing teams that are geographically dispersed.

Our future work will focus more on increasing the effectiveness of data collection and the creation of suitable APIs for the visualization of data that is collected in real time. Our future plans are to conduct similar observations in order to better understand the impact of sensors on participating GSD teams but also to assess how such data may help managing GSE projects.

Further work currently focuses on the use of other biometrics and physiological measurements such as heart rate monitors, eye trackers and electroencephalogram.

REFERENCES

performance. 32nd Annual Frontiers in Education. DOI: 10.1109/FIE.2002.1158709.


