A MECHANISM TO ASSESS THE RELATIONSHIP BETWEEN SOCIO-TECHNICAL CONGRUENCE AND PROJECT PERFORMANCE IN INCREMENTAL MODEL

W. A. W. M. Sobri¹, S. S. M. Fauzi¹*, M. H. N. M. Nasir², R. Ahmad² and A. J. Suali¹

¹Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 02600 Arau, Perlis, Malaysia
²Faculty of Computer Science and Information Technology, University of Malaya, Malaysia

Published online: 17 October 2017

ABSTRACT
Socio-technical congruence (STC) is one of the emerging areas in software engineering field. Prior study on STC suggests that the congruence between socio and technical dependencies leads to increase in task performance. However, little is discovered on how it relate to task performance in software development lifecycle particularly in incremental model. In this paper, we outline a method to investigate the relationship of STC and task performance in incremental model of software development life cycle.

Keywords: coordination; software development; software project; software engineering project; socio-technical congruence.

Author Correspondence, e-mail: shukorsanim@perlis.uitm.edu.my
doi: http://dx.doi.org/10.4314/jfas.v9i5s.6
1. INTRODUCTION
The competition to build a software product is getting fierce every day to provide products that are always relevant to current needs. First thing to remember, to ensure the success of a development project, coordination is essential for any developers involved [1-2]. Besides, developer which is coordinated through communication will determine the quality of software integration [1]. Previous research implies that the greater the geographic distribution, lessen the performance of the team [3]. Thus, under those circumstances, developer coordination is facilitated through activity such as email, formal meetings, instant messenger and other tools [4-6].
Prior study suggests applying Socio-Technical Congruence (STC) in software development helps to identify coordination needs in the early stage that leads to improve task performance [6]. However, despite the acknowledged impact of STC, there are still many fundamental questions that need to be addressed.
Study suggests different software development lifecycle has a different nature of development [7]. This may suggest the coordination needs in each software development life cycle is different. Incremental model especially as it is a model where the system is decomposed (increment) into parts, coordination between developers and the client is indeed vital to improving project productivity [8-9]. Little is discovered on how STC relates task performance in software development lifecycle, particularly in incremental model.
This paper aims to outline a method to investigate the relationship between STC and task performance in incremental model of software development lifecycle. This is in line with the role of software development lifecycle in each software engineering project as a catalyst to develop software systematically with lowering cost, time and resource.

2. LITERATURE REVIEW
Coordination is defined as when different people working on a common project and has come to mutual understanding and agreement. Coordination theory is a discipline of study derived as there is a need to coordinate activity of different actors in this recent years [10]. This is due to the fast change of global interdependencies which require organizations to prepare more
flexible method in order to manage activities [11]. To narrow the definition of coordination, it can also be defined as managing dependencies [8-9]. Obviously, coordination exist when there is interdependence among actors [12]. This lead to the need for developer to coordinate closely with whose technical dependencies affect his work in order to effectively develop software. Coordination in this study implies that performance of team member also affect others [13]. To elaborate, when there is less interaction between members their knowledge will be limited hence affect the amount of changes they are able to deal with.

Dependencies between components caused coordination problem among software engineering developers. In order to solve this problem, coordination mechanism is introduced where in this case developers need to do extra work by checking and take an action that ensures changes will not affect others component. Coordination mechanism varied based on coordination problem and certain dependencies [14]. Less study describes details about dependencies, hence it is hard to find what alternative processes help in certain situation. Novel theory called coordination theory introduced in 1994 to overcome this problem [15]. Based on prior study that uses coordination theory when observing companies, organizations which do the same task will do same related activities. As an example when bug fixing is assigned to an organization, then only related activity from fixing the bug until integrating into the system will be performed [16].

Coordination in distributed project is complex due to time, distance and cultural differences [17]. Rather than distributed, collocated project emphasize on communication between teams. As a matter of fact, communication is the main mechanism for teams with high coordination needs, for this reason, coordination failure also determines by communication [18,1]. In addition, shared work artefacts like code, bug, specification, has long been recognized as a medium that connects between developers through social networking web services at a distance [19]. Distributed projects stress on rigorous documentation, while encourages experimentation with communication tools [20]. Among the factors of software project failure if the teams do not communicate efficiently lead to suffering dysfunctional relationships [21]. Previously, it has been found that workers have been scheduling visits to the location of the other team as a way to improve coordination [20].
Coordination has two elements which are explicit and implicit. Explicit is a communication acts like email, chat, meetings or phone call while implicit is gaining information through listening and watching others conversations and work [22]. These are part of a factor contribute to strengthening the coordination, where trust is instilled among the team of software development via these elements [23]. Problem in coordination is found due to the intricate relationship, for instance complex and uncertainty of the interfaces and geographic dispersion. Thus, in order to gain coordination needed and sharing knowledge between the interdependent developer, there are ways to consider which is promote lateral communication and identification of dependencies which benefit for distributed project where they are notified of changes occur [24]. Prior study discovered that communication is in fact occurred notably in failed build where developers tend to correct technical dependencies, in the light of coordination and awareness is known to play an important role other than bridging the team [25]. However, awareness decreases with increasing distance between project member [25-26]. Among the challenges faced to maintain awareness in distributed teams is the lack of coordination and communication [26]. Thus, recent studies identify that simple method of communication is preferred to maintain awareness such as meetings, emails and mailing lists, instant messengers, phones and through bridging the coordination gap with experts [5]. Performance also decreases along with the increased of the gap in the distributed team, thus three solutions is suggested which is improving awareness of task dependency among developers, establish coordination through shared artefacts, and indirect communication through brokers [4]. Besides that, task performance will escalate in a structural congruence where formal team and communication path present [6]. Factors like team-building, effective communication, task and resource coordination, working harmoniously are crucial as an indicator towards team performance [27]. Communication among developers who are working on the modification of interdependent code helps prevent integration issue later on [28].

The traditional way of compromising with the complex project in order to minimize communication and coordination is by modularizing the system [29]. This is where the system is breaking down into a small module and is assigned to a team. On the other hand, Conway’s
Law conceptualizes that, the explanation of the collaboration between developers is that those who works on similar task must coordinate so that they are aware of the changes [29]. Moreover, through connecting domain experts with other developers, knowledge will grow wider [30].

Coordination will reduce when size and complexity of the project increases [18]. Furthermore, issues like location, time zone, and status negatively affected coordination [31]. This situation has become worse especially in the global software development due to the deficit of informal communication that leads to lack of awareness within team [32]. Though, now with the flooded variety of the latest technology and tools, communication and coordination activity still plays an important role for the developer in an organization [4]. Another factor to improve coordination, mainly in the global project is a manager’s contribution, where each of member duty needs to be informed early to avoid further misunderstanding and to be wise controlling conflict before it aggravates[23]. Other than that, creating and maintaining relationships between individuals on teams that coordinate is one of the ways to collaborate with other members [33].

Coordination through aligning between social and technical is considered to bring a positive effect on the task performance. Thus, IN [7] proposed STC as a measure which can calculate ‘fit’ between task-derived technical relationship and team social relationship. Socio as in socio-technical congruence refers to the relationships between developer, it is present when there is an interaction between developers involved in an organizational project [5]. Additionally, an interaction between team members is greatly helpful in generating mutual understanding of knowledge sharing [34]. Apart from that, communication has been a method that bridging the interaction between team members [1]. Whereas, technical is a work-derived relationship result from when two or more members from project team which has task dependencies is required to cooperate [5]. Technical components covered the processes, the tasks and the technology use in development project [35,6]. Initially, modular product structure has been the main method used by the organization in handling technical dependency [36]. Therefore, product is increment into smaller parts plus bound to technical dependency to reduce communication overhead [24]. The modularization theories state that
when technical interdependencies lessen within the module, then communication requirement within team members also lessen [37]. Hence, this will lead to increasing of integration issue with regard to the communication problem [38]. Customarily, technical information is shared through several ways such as forum, email and instant message among central and remote team in order to spread the knowledge [24].

Prior study shows every history of the bug are highly dependent on social, organizational and technical knowledge [39]. Many studies have shown an interest in social-technical aspects as a method in support software development, this recognition is due to an important of peer developers as well as other artefacts as knowledge resource [40]. In order to achieve successful collaboration between these two elements, socio and technical need to be congruence. Prior studies states STC as the match between technical relationship derived from technical dependencies and social relationship, whereby social relationship can be derived either from artefact or direct social relationship [25]. STC illustrate that productivity of software development project increase when coordination requirement align with relevant coordination activity. Coordination requirement in this study relates to the degree of coordination apply to the developers which run on changes, whereas actual coordination is described as coordination activity on which interaction occur among developers [6].

However, there is still additional study needed to investigate STC from different spectrum of software engineering project setting in order to strengthen the important role of STC in software engineering project.

Nevertheless, coordination in software project varied between different software development lifecycle (SDLC). Among the range of SDLC models, the most important and popular ones are the waterfall, spiral, agile, RAD, V and incremental model. Different SDLC models have their own way in developing software development projects, therefore coordination along the development period will differ between each SDLC model. Little is discovered on how development methodologies support coordination [41].

Hence, in order to further this study, among a varied range of SDLC, this research intends to look insight incremental model. Incremental model is a model where requirement are broken into parts and developed in each increment cycle where the first cycle starts with minimal part,
then the others part is added in the next cycles. Each requirement will go through complete cycle from gathering requirement, design and develop, until implementation [8]. Increment advantage is that in each cycle, client can provide feedback which help identify and resolve risk in advance [42]. Although the models have their own strengths and weaknesses, incremental model has the advantages of other models without their drawbacks[8]. For instance, risks are scattered between increments which makes it easier to overcome, this in contrast with, waterfall model which aggregates risk on its single iteration [43]. As well as spiral model, incremental model also emphasizes on risk management and process flexibility [8]. Moreover, increment enable developers to work parallel in cycles that overlap where this will minimize cost and time development risk [43]. Through mitigate early risk, project are able to be completed within schedule and lead to improve performance of the project. Prior study found that organization with high congruent is at advantage in task performance [4]. Nevertheless, less research studying the relationship between task performance and STC in software development lifecycle model. This research endeavors to investigate the relationship of STC in incremental model of software development lifecycle by using STC measure to identify task performance (resolution time). However, this paper will only highlight the method to investigate the relationship of STC and task performance in incremental model. The following section will discuss about the mechanism used to investigate how STC relate task performance in incremental model of software development lifecycle.

3. DESCRIPTION OF THE METHOD
In general, this section covers description of the methodology which includes settings and dataset, description of measure and congruence is explained. Apart from that, extraction and analysis method is briefly provided later.

3.1. Settings and Datasets
Using project which applying incremental model of software development lifecycle as their project guideline, help broaden research of STC in others software engineering context. The uses of incremental model are consistent with the project purpose to develop software in
more systematic and efficient way. Each increment in incremental model means one complete
development and feedback from the user in each increment has become an indicator whether
the right system is being built [8]. Feedback from users allows developer to be aware of
changes in the early stages and able to complete system within time and budget. Coordination
among developers in incremental model often occur as developers always need to discuss the
progress because this model allows frequent modification [9,45].
The nature of incremental model that started with minimal part until enormous and support
for large project is in fact, reflects open source software development. As it occur, open
source software often involve large-scale projects which are broken down by parts to the
members in different organization [44]. This is in line with this study which examine open
source project for analysis data. Table 1 describes the project selected.

<table>
<thead>
<tr>
<th>Table 1. Summary of the project</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Project Name</strong></td>
</tr>
<tr>
<td>CRUNCH</td>
</tr>
</tbody>
</table>

3.2. Description and Measure

Fig. 1 shows the basic theoretical model of STC:

![Proposed research model diagram](image)

3.2.1. Dependent Variable

Task performance is measured as the total time taken to complete the particular task
(resolution time) [46]. Whenever developers access and resolve particular change request,
apparently the time is recorded in modification request (MR) report. Hence, the resolution
time can be calculated based on that report.

3.2.2. Independent Variable

Actual coordination is calculated based on report in the mailing list which record communications that occur among developers involved [6]. Nonetheless, actual coordination for those developer without coordination need will not influence congruence [47]. Information related to developer interactions will be extracted from the project’s mailing list. Mailing list data will be extracted using .XML format. From the .XML format, developer interactions will be mapped into a table matrix. This will present interactions between developers. The matrix for developer interactions will develop using R script.

Coordination requirement infer task dependencies occurring in files-changed-together. Developers with the interdependence task need to coordinate, and accompanied by a proper communication will boost the task performance [6]. Files-changed-together is focus in this study as allow for multiple files involved to be sharing at the same time between developers with task dependencies. Shared files in this project acted as bridge between technical and social relationship, likewise developers who need to frequent communicate with the one sharing the dependencies.

Table matrices are developed to calculate coordination requirement. The Task Assignment matrix is developed to represent ‘people vs task’. Whereas, the Task Dependency matrix will be developed to represent ‘task vs task (files)’. Each matrix will be multiplied and will result in a people by people matrix, indicating the extent to which person i works on a task that share dependencies with the task done by person j. Defined below is the equation for coordination requirement matrix (denoted as CR), introduced by Cataldo and colleagues [6]:

\[
CR = TA \times TD \times TAT
\]

Based on the Equation (1), TA refers to the Task Assignment matrix. While TD refer to the Task Dependency matrix and TAT refer to the transpose of Task Assignment matrix. Information for coordination requirement will be extracted using an Excel format. From the Excel format, table matrices that provide information on the tasks assigned for each developer and files changed in each task will be developed. Each table will also be developed using R Script.
3.2.3. Congruence

Whereas, congruence is a result from comparing coordination requirement and actual coordination matrix. Empirical research suggest that when coordination requirement and actual coordination shows congruence reduced the development time [7]. Congruence equation is as follows, introduced Cataldo and colleagues [6]:

\[
\text{Diff}(CR,AC) = \text{card} \{\text{diffij} | CR_{ij} > 0 \& AC_{ij} > 0\}
\]

\[
|CR| = \text{card} \{CR_{ij} > 0\},
\]

\[
\text{Congruence} = \frac{\text{Diff}(CR,AC)}{|CR|}
\]

3.3. Extraction and Analysis Method

3.3.1. Extraction

Extraction data for this study is conducted through the Modification Request (MR) bug repository and also mailing list, where data is extracted using Mining Software Repository (MSR) techniques. MR bug repositories contain data involving file changes committed by developers who work on change request [48]. Description of data extraction and analysis method are included in Fig. 1.
3.3.2. Linear Regression Analysis

Linear regression analysis is applied to define the relationship among STC and task performance for every task in the project [48]. Since incremental model allow for overlapping between the cycle of increment, developers are able to coordinate their work in parallel in term of both communication and task [43]. Coordination between developers is required to resolve the error from prior cycle during the present increment. When social (communication) and technical (task) is in congruence, this will lead to improving their performance. This study will further the measure relationship between socio and technical which will result in development task performance. Correspondingly, study will demonstrate the relationship of STC in incremental model. Study using resolution time in order to measure the performance.
Equation to evaluate between those relationships is as follows:

\[ y = \beta_0 + \beta_1 \chi \]  

Based on equation stated, \( y \) represents task performance where \( \chi \) is the congruence. \( \beta_0 \) refer to the intercept and finally \( \beta_1 \) which act as gradient in that equation. By applying linear regression analysis, prior study found that congruence has significant effect on task performance. For instance, when geographical congruence, resolution time will decrease as the congruence rise [6].

4. CONCLUSION

The research is helpful by providing information about the mechanism on how to investigate relationship between STC and task performance in incremental model of software development lifecycle. Empirical evidence argues that by a method of breakdown task help reduces dependencies which yield good performance [49]. This is in line with the incremental model itself that emphasizes the development based on task partition. This study then could be used for future reference for researchers that study related to this topic.

5. ACKNOWLEDGEMENTS

We gratefully acknowledge support by Ministry of Education (MOE), Malaysia and UniversitiTeknologi MARA, Malaysia under Research Acculturation Collaborative Effort under Grants No 600-RMI/RACE 16/6/2 (10/2014).

6. REFERENCES


[14] Crowston K. A coordination theory approach to organizational process


2007, pp. 81-90


[40] Nakakoji K, Ye Y, Yamamoto Y. Comparison of coordination communication and expertise communication in software development: Motives, characteristics, and needs. In International Symposium on Artificial Intelligence, 2009, pp. 147-155


