

TASK ALLOCATION IN A GLOBALLY DISTRIBUTED SOFTWARE DEVELOPMENT ENVIRONMENT

Chintan Amrit & Jos van Hillegersberg

Department of IS&CM, University of Twente, The Netherlands

ABSTRACT

In this paper we tackle one crucial part of the coordination problem in globally distributed software development, namely the problem of task assignment in a team. There are a number of ways to split work among the globally distributed sites. The work could be split according to the process state it is in, according to the product structure or even according to the product features. These different strategies could be more or less effective depending on a variety of circumstances. We propose a technique to test for problems in coordination through allocation of tasks, by observing the task allocation and advice network structures.

We perform a pilot case study based on 4 teams of Masters Student working in a globally distributed environment (India and The Netherlands) where we analyse the social network structures along with the task distribution in each of the teams. By calculating metrics on these networks we represent a technique by which one can locate coordination problems in distributed software teams.

Keywords:

Software Development, Coordination, Social Network Analysis, Teams, and Task Allocation.

1. INTRODUCTION

Large scale software development is an inherently collaborative, team based process, and hence requires coordination and control in order to make it successful. The advent of global distribution of software development has made this collaboration more difficult due to various reasons including physical distance, differences in time, cultural differences etc (Herbsleb and Moitra 2001). Although research on global software development argues the use of communication technologies to alleviate problems caused by separation of workers in time and space, studies have often found them to be not as effective as publicized (Bal and Foster 2000). A nearly total absence of informal communication among developers makes coordinating work in a globally distributed setting more critical (Herbsleb and Grinter 1999). Traditionally, most project management approaches for improving coordination in software development have emphasized on one of the following three methods of technical innovations:

1. Development of new and enhanced methods and tools (Kraut and Streeter 1995; Grinter 1998; Herbsleb, Mockus et al. 2001).
2. Modularization both technical (Object Oriented Programming) or managerial such as the organizational separation of requirements, design concerns, coding and testing functions to encapsulate the behavior of program elements and individual software professionals, and thereby reduce the needs for coordination (Kraut and Streeter 1995).
3. Formal procedures, both technical, such as version control software, case tools, and specification languages such as test plans, delivery schedules and requirements documents to control communication among development personnel (Kraut and Streeter 1995; Herbsleb, Mockus et al. 2001).

While these techniques contributed to a modest increase in software productivity over the past twenty years, they only partially address the problem of coordination (Kraut and Streeter 1995). A more recent approach has involved improved project management practices applied to software development process. These practices focus on improving task decomposition, task assignment and work group coordination, which are considered important issues in the context of Coordination (Kirsch 1996; Crowston 1997; Andres and Zmud 2001). In dealing with coordination as result of assigning tasks it is noted that tasks that have low coupling can be more easily coordinated in a distributed team which has a high degree of cohesion (Olson and Olson 2000; Sakthivel 2005). Although loose coupling across sites holds promise as a structural solution for distributed teams, it may not be a panacea. Although teams may not have day-to-day requirements to interact, loose-coupling can mask true interdependencies and can lead to more severe problems at the point of handoff (Grinter, Herbsleb et al. 1999). An alternate method of allocating tasks across

sites is to split the product according to its structure and develop the different product parts at different sites, resulting in an organizational structure that mirrors the product structure (Conway 1968; Herbsleb and Grinter 1999).

In this paper we show a methodology of testing whether a particular way of splitting tasks is indeed beneficial to the team. We use this technique to test if this method of allocating tasks is indeed a solution to determine problems in coordination. The technique we employ is to test a metric indicating a relation between group cohesion and task coupling with a pilot case. The pilot survey we conducted was on four teams of students who worked on a similar software design assignment. We then observed the social network of the teams concerned, along with the distribution of the software design tasks among the team members. We then compared the performance of the team members depending on the degree of cohesion of their advice and task networks related to their performance to come up with insights on whether problems in coordination related to assignment of tasks can be detected.

The rest of the paper is structured as follows; section 2 deals with an overview of social network techniques applied to deal with coordination, section 3 contains an overview of the literature dealing with formation of teams and allocation of tasks among them, section 4 deals with the details of the empirical test we performed on the four distributed teams of students, section 5 contains a discussion of the results, section 6 deals with the limitations of this particular case study and finally section 7 summarizes the results of this study and mentions some recommendations for further research.

2. SOCIAL NETWORK METRICS

The problem of coordination can be better explained by first illustrating the concepts of social networks and centrality measures. A social network consists of a set of actors (“nodes”) and the relations (“ties” and “edges”) between these actors (Wasserman and Faust 1994).

Researchers have developed a variety of metrics for quantifying the differences in network structure. Among the many frequently used metrics at the actor level are degree centrality (the extent to which actors send or receive ties) and *betweenness centrality* (the extent to which actors have ties with others who are not directly connected). Metrics used to describe networks include: Density (the ratio of the pairs of nodes that are mutually reachable to the total number of pairs of nodes) and Centralization (Difference between the centrality scores of the most central actor and those of all other actors in a network is calculated, and used to form the ratio of the actual sum of the differences to the maximum sum of the differences).

How is a group defined, from the network perspective? The construct of a group, when used in the social network literature has had two primary meanings: (a) a structural feature of a network, or (b) an exogenously determined or imposed category. According to the first meaning, groups (cliques, a maximally complete subgraph) are subsets of fully connected, or almost fully connected, nodes within some population (Katz, Lazer et al. 2003).

The problem of task allocation among team members is closely related to the network structures formed between them. Several researchers have asked the question “What is the optimal network for group performance?” Many have broadened the scope of investigation by moving from the laboratory to the field.

Sparrowe, Liden, Wayne and Kraimer (Sparrowe, Liden et al. 2001) demonstrated a relation between network structure and both in-role and extra-role performance in a field setting. They replicated earlier findings; with a complex task, and found that groups with decentralized communication patterns perform better than groups with centralized communication patterns. Cummings and Cross (Cummings and Cross 2003) also found that groups with decentralized communication patterns outperformed more centralized groups.

Other researchers have focussed on the number rather than the pattern of communication links among group members. Baldwin, Bedell and Johnson (Baldwin, Bedell et al. 1997) and Reagans and Zuckerman (Reagans and Zuckerman 2001) established in field studies that groups with more ties performed better than groups with fewer ties.

3. TEAM FORMATIONS AND TASK ALLOCATION

Teams are the basic building block for many contemporary business organizations. We focus on how we can improve coordination in software development projects using the concepts of coordination between and among teams keeping task assignment as a moderating variable. Coordination refers to team-situated interactions aimed at

managing resources and expertise dependencies (Faraj and Sproull 2000). Research on software development teams has found that team performance is linked with the effectiveness of teamwork coordination (Kraut and Streeter 1995).

Faraj and Sproull (Faraj and Sproull 2000) take two perspectives on coordination: administrative coordination and expertise coordination. They claim that administrative coordination (management of tangible and economic resource dependencies) is good for simple routine tasks, while for complex non-routine intellectual tasks, expertise coordination (the management of knowledge and skill dependencies) become more important. Through expertise coordination the team can recognize and access expertise when it's needed.

Stewart and Barrick (Stewart and Barrick 2000) build on organization-level findings and show that differences in how responsibilities are apportioned and coordinated correspond to variance in performance at the team level. They also show that the effect of these social elements is moderated by technical demands (tasks), consistent with socio-technical systems theory.

Sparrowe et al. (Sparrowe, Liden et al. 2001) hypothesize that centrality in a work group's advice network will be positively related to an individual's job performance. Where centrality in the advice network reflects an individual's involvement in exchanging assistance with co-workers and engaging in mutual problem solving. An individual who is central in the advice network is, over time, able to accumulate knowledge about task-related problems and workable solutions (Baldwin, Bedell et al. 1997). While the central individual develops problem solving capability and serves as a valued resource for future exchanges with co-workers, those individuals who are in peripheral positions in the advice network find it difficult to develop expertise and competencies for high levels of performance (Sparrowe, Liden et al. 2001). Hence, Sparrowe et al. (2001) hypothesize that centralization in a work group's advice network is negatively related to group performance.

Yang and Tang (Yang and Tang 2004) try to analyse the relation between team structure and ISD performance using a social network approach. They show how the structural properties of the work groups fluctuate during the various phases of Software Development, and how group cohesion and centrality are related to the final ISD performance. Though Yang and Tang (2004) do show how social research methods can be used to tackle "group process" factors, they do not deal with task allocation nor do they illustrate how one can solve the problem of task allocation among team members.

Sakthivel S. (Sakthivel 2005) describes the Level of Virtual Work as the feasibility of offshore development of the task depending on the degree of task coupling and the degree of cohesion among the team members. According to him group work that has a high task coupling for a group with a low degree of cohesion must be done in a face to face environment, while group work that has a low degree of task coupling for a group with a high degree of cohesion can be done in a globally distributed environment (Sakthivel 2005).

Grinter and Herbsleb (1999) suggest the chief motivation for the assignment of tasks (involving a search for experts) in R&D projects to be:

Functional Area: an expertise of distant systems,

Product Structure: an understanding of the internals of components built at remote sites,

Process: knowledge about what happens during other processes,

Customisation: knowledge of core or the knowledge of how the core is customized depending on which site is involved.

Yet another strategy divides the product according to its structure and develops various product parts at different sites, resulting in an organizational structure that mirrors the product structure (Conway 1968; Herbsleb and Grinter 1999). In splitting work along the lines of product structure one must consider the modular design of the product in order to isolate the effect of changes (Parnas 1972). MacCormack and colleagues (MacCormack, Rusnak et al. 2006) reiterate Conway's argument (Conway 1968) when they compare commercial and open source development. As the software developers were collocated in the commercial project, it was easier to build tight connections between the software components, therefore producing a system more coupled compared to the similar open source project with distributed developers.

While the Conway's Law relation between the task and coordination of the developers has been validated by several

empirical studies (Curtis, Krasner et al. 1988; Grinter, Herbsleb et al. 1999; Herbsleb and Grinter 1999; Sosa, Eppinger et al. 2004), none of these studies have used the social network perspective.

In this paper we test Conway's Law relation between group cohesion and task coupling in social networks with a pilot case and see if we can determine the quality of task splitting between the distributed teams. In our pilot case we assign a similar task structure to four teams and see how their performance is affected by the structure of their advice and task networks.

4. EMPIRICAL TEST

4.1 PROJECT QUESTIONNAIRE

A pilot survey was conducted on 4 globally distributed teams of Masters Students consisting of 8, 8, 7, 7 students respectively. Approximately half of the members of each team consisted of Dutch students located in a Dutch university and other half were Indian from a university located in India. The students were asked to select a topic for a design-based project, and complete four design tasks involving the creation of a vision document, activity diagram, use case and class diagram for the selected project topic. The data was collected with the help of a questionnaire, in which among other questions we asked:

Rate your contribution (relative to the average team member) in creating the Vision Document?

Rate your contribution (relative to the average team member) in creating the activity diagram?

Rate your contribution (relative to the average team member) in creating the Use Case diagram?

Rate your contribution (relative to the average team member) in creating the class diagram?

Mark your team members from whom you regularly sought information and advice to help in your project work.

Mark your team members, whose advice you did not seek, during the course of the project.

Mark your team members who in your opinion are very dependable in executing a crucial part of the project.

The first 4 questions had options out of a scale of 5, while the questions 5 to 7 had the list of all the team members to choose from (multiple members could be checked for each answer). From the answers to questions 1 through 4 the relative contributions to the particular task were obtained. This helped in drawing the 2-mode task network. Where we can see the team members assigned to the tasks (making the Use case diagram, activity diagram etc.) with the links having a particular weight corresponding to the amount of effort each team member has put in performing the task. The contributions ranged from 1 (no relative contribution) to 5 (full contribution). The contribution was based on team member perception, so it was possible for all the members of the team to think that they had done the task themselves (all of them to fill 5). From the answer to questions 5 and 6 the advice network of the group members was obtained. The answer to question 6 confirmed the network obtained from the answer to question 5. We used the data from the advice and task networks to calculate our metrics in this paper.

4.2 MEASURES

4.2.1 Network density

In binary network data, density is the proportion of actual nominations among the total possible number of nominations (Wasserman and Faust 1994). This was computed by using the density function of UCINET for networks, using total number of ties present divided by the total number of all possible ties. The Density of the Advice Network (DAN) is the network density taken for the advice network, while the Density of the Task Network (DTN) is the network density taken for the task network.

4.2.2 Task Density

The task network is a 2-mode network (Borgatti and Everett 1997). In order to find the density the weighted task network was first dichotomised using the standard dichotomise routine in UCINET software package. The cut-off value was considered to be 3, as the contribution of less than 3 on 5 was considered negligible (also because 3 is the

median on a scale of 1 to 5). So the dichotomization rule was as follows:

$y(i, j) = 1$ if $x(i, j) \geq 3$, and 0 otherwise.

4.2.3 Team Performance

The final performance of the team was rated according to the following metrics:

Time taken for project completion

Documentation and its revision history

Quality of deliverables

Relevance of alternative solutions suggested

The final grade scored by the individual teams was considered (out of 10) (Table 2).

Team	Density of Advice Network (DAN)	Density of task network (DTN)	DTN/DAN
Team 1	0.3929	0.7188	1.824
Team 2	0.3036	0.8125	2.676
Team 3	0.5000	0.9643	1.9286
Team 4	0.4048	0.8571	2.1173

Table 1: Statistics of some of the network measures

5. DISCUSSION AND RESULTS

Conway (Conway 1968) found a very close relation between the structure of a system and the structure of the organization which designed it, in other words he found that there is a *homomorphism* from the graph of the designed system to the graph of the organization. Subsequently this pattern of task allocation has got to be called Conway's Law (Herbsleb and Grinter 1999; Herbsleb and Grinter 1999). In our empirical study we provide a technique using social networks to test whether the task allocation among globally distributed teams as suggested by Conway's Law is good for the project. To begin with we assigned a similar task structure for each of the four distributed teams. We then constructed the *Coordination Requirement* network (Cataldo, Wagstrom et al. 2006) based on the task networks of the four teams. The *Coordination Requirement* network is a network where all the people working on dependent tasks or on the same task are linked to each other (Cataldo, Wagstrom et al. 2006). The *Coordination Requirement* Network for each of the four teams resulted in a complete graph (a graph with edges linking every pair of nodes). This implied that the team members, in each of the four teams were all dependent on each other. So in order to measure the amount by which each of the teams obeyed the pattern suggested by Conway, we simply had to check how *homomorphic* the Advice and Task Networks were. To calculate this we can divide the Density of the Task Network with the Density of the Advice Network and the closer this number (DTN/DAN) is to one, the more *homomorphic* the networks are and the better the particular team obeys the pattern suggested by Conway's Law. Though this ratio (DTN/DAN) doesn't always indicate homomorphism, in this particular case, where each of the *Coordination Requirement* networks is a complete Graph (each of the Task Network is very dense) for all the four teams, this ratio shows the extent of network homomorphism. The result of the calculation is shown in Table 1.

We expect a team with a higher density in the Advice network to perform better than a team with a lower density (Sparrowe, Liden et al. 2001). Thus, according to the density of the advice network we expect the teams to have performed in the order: Team 3 > Team 4 > Team 1 > Team 2. According to the ratio DTN/DAN we expect the team

whose ratio is closer to one to perform better than the team whose ratio is further from 1, as the team whose ratio is closer to one would have their communication network similar to their task dependencies and hence obey Conway's Law. On calculating the ratio (Table 1.) we arrive at the rating: Team 1 > Team 3 > Team 4 > Team 2, as the ratio for Team 1 is the closest to one and for Team 2 the furthest.

We obtained the final ranking by evaluating the quality of the deliverables, which was: Team 3 > Team 1 > Team 4 > Team 2 (Table 2). Though this result is different from the one predicted by Conway's Law we can perhaps explain the result by examining the advice networks of Team 1 and Team 3. In all both the figures 1 and 2 we see that the advice networks consists of a *structural hole* (Burt 1992) on either side of which represents the Indian component (a connected subgraph) and the Dutch component (another connected subgraph) of the globally distributed team. The structure of these advice networks further shows the presence of gate keepers, who are persons involved in communicating with their global (Dutch or Indian as the case maybe) counterparts. In the case when one of the components (Indian or Dutch) in the advice networks is not well connected we see that the performance of the groups suffers. This is evident in figure 1, where the Indian component is sparsely connected. This can be further be used to show the differences in predicting the performance of Team 1 and Team 3. Though the differences can be due to several factors, as is common in analysing most Software Development Projects (Kraut and Streeter 1995), from a network perspective we can say that the visible difference in their advice networks might have contributed to their difference in performance. We find that the advice network of the Indian component in Team 1 (figure 1) is not so well connected as the advice network of the Indian component in Team 3 (figure 2). From this difference in structure one can say that the lack of a well-connected advice network of a component of a globally distributed team might have had an impact on the performance of the team, causing Team 1 to perform worse than Team 3.

6. LIMITATIONS

The primary limitation is that the empirical data is not very strong, as the sample size is too small for any kind of statistical analysis. The idea behind this paper is to illustrate how this methodology can be used for future analysis on larger samples.

The students in the Dutch University were without much industrial experience while the students in the Indian University had some industrial experience. Further, the teams were composed of relative equals and had limited existence, making them unlike most real world teams. In future research, more case studies/surveys need to be conducted on people working in the software industry.

The tasks, which were a part of their course project, were small design tasks of software development. Also, the number of modules was limited to four, while in a normal software development project the number of modules runs into thousands and corresponding number of tasks per person can be much higher. So, in future such a study can be conducted in an IT company with more demanding project requirements.

Team	Grades	Evaluation of Quality	Overall Rank
Team 1	7.5	10	2
Team 2	7	8	4
Team 3	7.5	15	1
Team 4	7	10	3

Table 2: The result of evaluation of the Teams

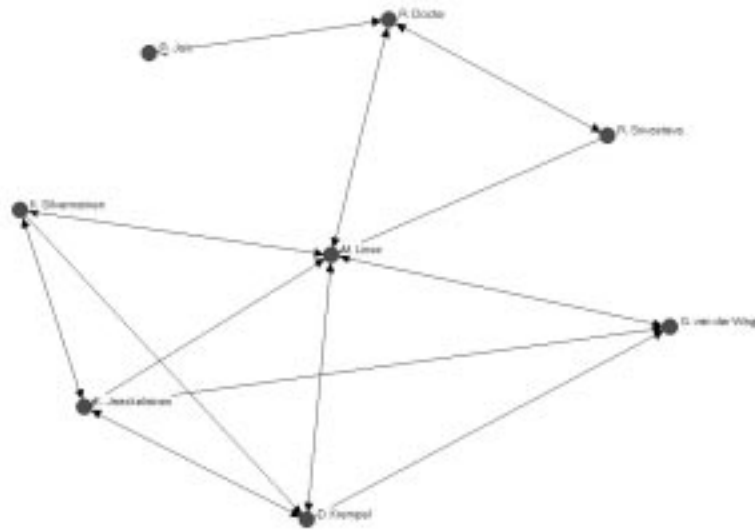


Figure 1: Advice Network of Team 1

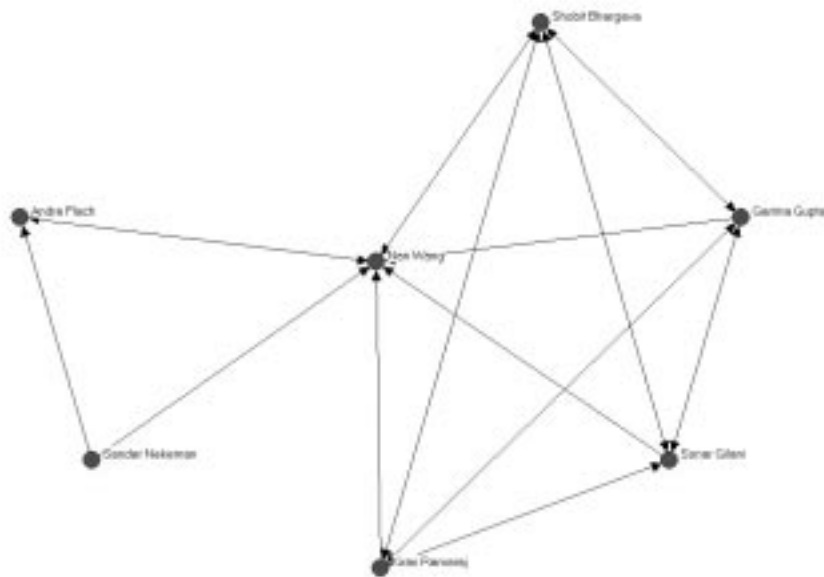


Figure 2: Advice Network of Team 3

7. CONCLUSION AND PRACTICAL IMPLICATIONS

In this paper we have come up with insights on the methodology by which one can use social network analysis to improve the coordination in the globally distributed software development process of an IT company. This study adds to the growing body of knowledge in managing distributed software development. We have shown a technique by which we can determine how well the tasks have been assigned to distributed teams. By calculating the density

of the advice network and the task network we have shown how far the coordination dependencies in the task network are satisfied by the communication in the advice network and hence how far the team's task structure satisfies Conway's Law. Though many previous empirical studies have shown this in the past (Grinter, Herbsleb et al. 1999; Herbsleb and Grinter 1999; Sosa, Eppinger et al. 2004), we show how one can determine which method of splitting work or task allocation among the distributed teams is more beneficial, through the use of social networks. In other words the more a teams' task and advice networks obey Conway's Law (are Homomorphic) the better would be the coordination among the team members which would positively impact the performance of the teams.

In future research, more surveys/case studies can be conducted on larger groups in the Software Industry to test whether the technique presented through using social networks is indeed useful. Further, such a test can be conducted in a longitudinal manner at different stages of the software development project. Adding the time component can throw more light on the way social networks change and develop during the course of a project. It could also be used to understand the different task allocation strategies needed to make a project successful.

We have developed a software tool called *TESNA* (TEchnical and Social Network Analysis) which can gather data on the technical and social networks automatically and measure metrics of both software and social networks over a period of time. In future research we plan to test this tool on different case studies to study the task allocation structures in large software development organizations. We have also begun to investigate open source software development, to see the differences between corporate and Free/Libre open source task structures.

ACKNOWLEDGEMENTS

We would like to thank Prof. V. Sridhar of MDI Gurgaon for helping us design the questionnaire. We would also like to thank the students in the Information Management course in Rotterdam School of Management as well as the students of MDI Gurgaon

REFERENCES

- Andres, H. P. and R. W. Zmud (2001). "A Contingency Approach to Software Project Coordination." *Journal of Management Information Systems* Vol. 18(Issue 3): p41.
- Bal, J. and P. Foster (2000). "Managing the virtual team and controlling effectiveness." *International Journal of Production Research* 38: 4019-4032.
- Baldwin, T. T., M. D. Bedell, et al. (1997). "The Social Fabric of a Team-Based M.B.A. Program: Network Effects on Student Satisfaction and Performance." *The Academy of Management Journal* 40(6): 1369-1397.
- Borgatti, S. P. and M. G. Everett (1997). "Network analysis of 2-mode data." *Social Networks* 19(3): 243-269.
- Burt, R. (1992). *Structural Holes*. Cambridge, MA, Harvard University Press.
- Cataldo, M., P. Wagstrom, et al. (2006). Identification of coordination requirements: implications for the Design of collaboration and awareness tools. *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*. Banff, Alberta, Canada, ACM Press.
- Conway, M. (1968). How do Committees Invent. *Datamation*. 14: 28-31.
- Crowston, K. (1997). "A Coordination Theory Approach to Organizational Process Design." *Organization Science* 8(2): 157-175.
- Cummings, J. N. and R. Cross (2003). "Structural properties of work groups and their consequences for performance." *Social Networks* 25: 197-210.
- Curtis, B., H. Krasner, et al. (1988). "A Field-Study of the Software-Design Process for Large Systems." *Communications of the Acm* 31(11): 1268-1287.
- Faraj, S. and L. Sproull (2000). "Coordinating Expertise in Software Development Teams." *Management Science* 46(12): 1554-1568.
- Grinter, R., E. (1998). Recomposition: putting it all back together again. *CSCW '98: Proceedings of the 1998 ACM conference on Computer supported cooperative work*. New York, NY, USA: 393—402.
- Grinter, R., E., J. D. Herbsleb, et al. (1999). The geography of coordination: dealing with distance in R&D work. *Proceedings of the international ACM SIGGROUP conference on Supporting group work*. Phoenix, Arizona, United States, ACM Press.
- Herbsleb, J., D. and R. Grinter, E. (1999). Architectures, Coordination, and Distance: Conway's Law and Beyond. *IEEE Softw.* Los Alamitos, CA, USA. 16: 63—70.
- Herbsleb, J., D. and R. Grinter, E. (1999). Splitting the organization and integrating the code: Conway's law revisited. *ICSE '99: Proceedings of the 21st international conference on Software engineering*. Los Alamitos, CA, USA: 85—95.
- Herbsleb, J., D., A. Mockus, et al. (2001). An Empirical Study of Global Software Development: Distance and Speed. *ICSE*

proceedings. Los Alamitos, CA, USA. **00**: 0081.

Herbsleb, J. D. and D. Moitra (2001). "Global software development." Software, IEEE **18**(2): 16-20.

Katz, N., D. Lazer, et al. (2003). "Network theory and small groups." Small Group Research **35**(3): 307-332.

Kirsch, L. J. (1996). "The Management of Complex Tasks in Organizations: Controlling the Systems Development Process." Organization Science **7**(1): 1-21.

Kraut, R., E. and L. Streeter, A. (1995). Coordination in software development. Commun. ACM. New York, NY, USA. **38**: 69—81.

MacCormack, A., J. Rusnak, et al. (2006). "Exploring the structure of complex software designs: An empirical study of open source and proprietary code." Management Science **52**(7): 1015-1030.

Olson, G. M. and J. S. Olson (2000). "Distance matters." Human-Computer Interaction **15**(2-3): 139-178.

Parnas, D. L. (1972). On the criteria to be used in decomposing systems into modules. Commun. ACM. New York, NY, USA. **15**: 1053—1058.

Reagans, R. and E. W. Zuckerman (2001). "Networks, Diversity, and Productivity: The Social Capital of Corporate R&D Teams." Organization Science **12**(4): 502-517.

Sakthivel, S. (2005). "Virtual workgroups in offshore systems development." Information and Software Technology **47**(5): 305-318.

Sosa, M. E., S. D. Eppinger, et al. (2004). "The Misalignment of Product Architecture and Organizational Structure in Complex Product Development." J Manage. Sci. **50**(12): 1674-1689.

Sparrowe, R. T., R. C. Liden, et al. (2001). "Social networks and the performance of individuals and groups." Academy of Management Journal **44**(2): 316-325.

Stewart, G. L. and M. R. Barrick (2000). "Team Structure and Performance: Assessing the Mediating Role of Intra-team Process and the Moderating Role of Task Type." The Academy of Management Journal **43**(2): 135-148.

Wasserman, S. and K. Faust (1994). Social Network Analysis: methods and applications, Cambridge University Press.

Yang, H.-L. and J.-H. Tang (2004). Team structure and team performance in IS development: a social network perspective. Information and Management. Amsterdam, The Netherlands, The Netherlands. **41**: 335—349.

