COalitions in COOperation Networks (COCOON)

Social Network Analysis and Game Theory to Enhance Cooperation Networks

by
RORY SIE

COALITIONS IN COOPERATION NETWORKS (COCOON)

Social Network Analysis and Game Theory to Enhance Cooperation Networks

The research reported in this thesis was carried out at the Open University of the Netherlands in the Centre for Learning Sciences and Technologies and was partly carried out within the IdSpace project. The idSpace project is partially supported/co-funded by the European Union under the Information and Communication Technologies (ICT) theme of the 7th Framework Programme for R&D. This document does not represent the opinion of the European Union, and the European Union is not responsible for any use that might be made of its content.

SIKS Dissertation Series No. 2012-33
The research reported in this thesis was carried out under the auspices of SIKS, the Dutch Research School for Information and Knowledge Systems.

ISBN/EAN 9789491465680

Cover design: Bert Uytley
Printed by: Datawyse, Maastricht, The Netherlands

COALITIONS IN COOPERATION NETWORKS (COCOON)

Social Network Analysis and Game Theory to Enhance Cooperation Networks

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Open Universiteit
op gezag van de rector magnificus
prof. mr. A. Oskamp
ten overstaan van een door het
College voor promoties ingestelde commissie
in het openbaar te verdedigen

op vrijdag 28 september 2012 te Heerlen
om 13.30 uur precies

door

Rory Liang Lee Sie
geboren op 6 juli 1982 te Amsterdam

Promotor
Prof. dr. P.B. Sloep
Open Universiteit

Co-promotor
Dr. M.E. Bitter-Rijpkema
Open Universiteit

Overige leden beoordelingsommissie
Prof. dr. D. Griffiths
University of Bolton, United Kingdom

Prof. dr. E.O. Postma
Universiteit van Tilburg

Prof. dr. M.C.J. Caniëls
Open Universiteit

Dr. V.G. Dimitrova
University of Leeds, United Kingdom

Dr. P.M. van Rosmalen
Open Universiteit

CONTENTS

CHAPTER 1

General Introduction

In this introductory chapter we introduce the notions of a cooperation network and some of its siblings, such as innovation networks, research networks and learning networks. It is these networks that our research focuses on and we discuss the questions and hypotheses that we investigate with respect to them. This includes inventorying what we already know about such networks from the extant literature.

1.1 Warm-up

Once every four years the Fédération Internationale de Football Association (FIFA) organizes the World Cup Football. Every World Cup takes places in a different country, for example Uruguay (1930), England (1966) and South Africa (2010). Sometimes, the event is even organized by two countries, such as South Korea and Japan in 2002. Traditionally, at every World Cup a new football is introduced[1]. Every new ball has new, innovative characteristics such as better accuracy or better ball control. For instance, the 'Jabulani' - the official World Cup ball introduced at the South Africa World Cup - exhibits improved stability during flight, due to its *Aero grooves* (Adidas, 2009). New balls are intensely tested by both professional football players and the FIFA. The FIFA determined a number of test criteria such as perfect roundness, flight characteristics and absorption of water.

Now, let's imagine Adidas appointed you as their new football engineer. You have to design a ball that is an innovation relative to the ball used at the previous World Cup. In the past, you have worked in architecture, so you are familiar with some surface technology, but football engineering is a 'whole new ball game' to you. You have to meet the standards set by the FIFA, and you need to satisfy your customer, the professional football players. Furthermore, the ball will not only be used during the World Cup, but will also be sold in stores for the public. As football seems to be played by each and every social class, the ball needs to be affordable. Thus, there are a lot of constraints and criteria, but this also provides an opportunity for you to excel at your job.

Just to acquaint what Adidas came up with so far, you start examining the characteristics of the previous balls they created. You do not want to disturb the Adidas management with a ball that contains old technology. You start summing up the advantages and disadvantages of the current balls. You may even want to ask professional soccer players what they value in a good soccer ball. Some players may mention good grip, because they want to control the ball under any weather circumstance. Players that are specialised in taking direct free kicks on goal may find it important that a ball can curve around a defensive wall of players.

Since 2005, Adidas has been working on *Goal Line Technology*. They have created several types of goal line monitoring devices, including technology inside soccer balls that transmits signals allowing one to detect whether or not a ball has crossed the goal line. Another example of goal line technology is the use of cameras for that purpose. Yet, the technologies have not yet shown to be reliable in one hundred percent of the cases. Referees are not one hundred percent reliable either, but

[1] Actually, this has been a tradition since the 1970s when Adidas developed the Telstar for the World Cup in Mexico.

they are human. For technology to be implemented and used alongside the referee, the FIFA wants one hundred percent reliability, or at least something very close to that.

Suddenly, you are struck by this wonderful idea to put light-sensitive nanotechnology onto the surface of the ball. Light sensors can already be used to distinguish body positions, such as standing, sitting and running (Maurer, Smailagic, Siewiorek, & Deisher, 2006). A white goal line reflects more light than the surrounding green grass, so it should be possible for the ball to 'see' where it is. Together with other, existing technology such as goal line cameras, this may be the missing piece of the puzzle that will perfect goal line technology.

Knowing that you are not an expert on all areas of football design, you start looking for experts that can help you design this ball. Since you are new to this working area, you do not personally know anyone. How could you know who are the experts in distinct soccer ball technology areas such as aerodynamics, surface technology, water absorption and testing? In other words, you lack a certain degree of *awareness* of who is an expert, and on what topic.

Alternatively, let us assume that you are not entirely new to this field, and you know all experts worth knowing. Who will guarantee that you pick the *right* experts to work together with? You have to form the right design team, that is, a team that is able to collaborate without too many interpersonal problems. Every individual is unique, and forming a team of unique individuals inherently poses the threat of whether these personalities and behaviours are compatible. For instance, research in the USA has shown that there is an inverse relationship between racially diverse teams and in-group support (Bacharach, 2005).

Furthermore, the team should reflect the knowledge that is needed to design the new soccer ball you had in mind. In order to innovate and improve the balls that are already on the market, you need to include top professionals in your soccer ball design team. Assuming that you know people who are the acknowledged experts in this domain, there still are numerous other problems and considerations that you have to take into account. How do you know they indeed are experts? Will they be willing to work together? Are they available at the right time and location? Will their personalities match? Intuitively, we can tell that if two people have a mismatch in personality, they are not likely to work together smoothly.

The above example shows that innovation is sensitive to several factors that influence both the cooperation process and the decision whom to cooperate with. A key assumption of this thesis is that innovation networks are a type of cooperation networks, and that they share a lot of characteristics with other social networks in which we cooperate, such as research networks for doing collaborative research and learning networks for knowledge sharing and creation. It is cooperation networks that we are interested in in this thesis, primarily in the form of their manifestations as innovation networks, research networks and learning networks.

1.2 Cooperation networks

We define a *cooperation network* as a network of actors that have the intention to work together. The nodes in the network typically represent human beings, and the edges between these nodes represent their shared cooperative intentions. For the purpose of this thesis, cooperation presupposes the cooperators' intention of going into the same direction (coordination), but does not necessarily require the same goal.

Cooperation can be illustrated by a famous story that Mary Parker Follett (Follet & Metcalf, 2003) once told about two sisters that fought over a single orange. They had dissimilar goals: One sister wanted the orange to make juice, the other sister wanted the peel to bake a cake. They made a compromise by splitting the orange in half, whereas they could have kept their distinct goals: one would get all the juice, and the other would get all the peel. The example of the sisters and the orange explicates the difference between cooperation and collaboration. Cooperation requires two individuals to share intentions, but their individual goals remain the same (make juice and bake a cake). Collaboration requires two individuals to share intentions and have a common goal (share the orange) without taking into account the individual goals. In this case, the sisters could have optimized their outcome by keeping their distinct goals and cooperate.

The story we sketched above is an example of how innovation using a social network typically occurs. We search our network for people that are knowledgeable, know where to get the knowledge (so-called *knowledge brokers*), or people that can help us get our ideas accepted. If we use our social network to enhance the innovative process, we call this *networked innovation* (Swan & Scarborough, 2005). Innovation networks - the networks in which we perform networked innovation - are a type of cooperation network. In an innovation network, individuals share the intention to innovate, but they may have different goals. Similarly, we have learning networks in which we intend to learn (Sloep & Berlanga, 2011), and research networks in which we intend to perform research (Reinhardt, 2012).

This thesis focuses on how we can assist cooperation in such networks. Obviously, assisting in cooperation is easier said then done. Quite in general, before assistive tools and procedures can be developed, it is necessary to have a thorough understanding of what might hamper cooperation. This is what we will now turn our attention to.

1.3 Common problems in cooperation networks

We distinguish four types of problems (Figure 1.1). First, we have *intrapersonal problems*; problems that influence the individual when engaging in cooperation through a social network. These problems may involve cognitive problems such as lack of awareness, bounded rationality, information overload (see below for their explanations). Second, we have *interpersonal problems*; problems that influence the relationship between two individuals, such as knowledge sharing problems. Third, we have *procedural or structural problems*; constraints that are put on us while we are cooperating. Finally, we have *exogenous problems*; factors that lie beyond the control of the individuals that are cooperating, such as time, money, and culture.

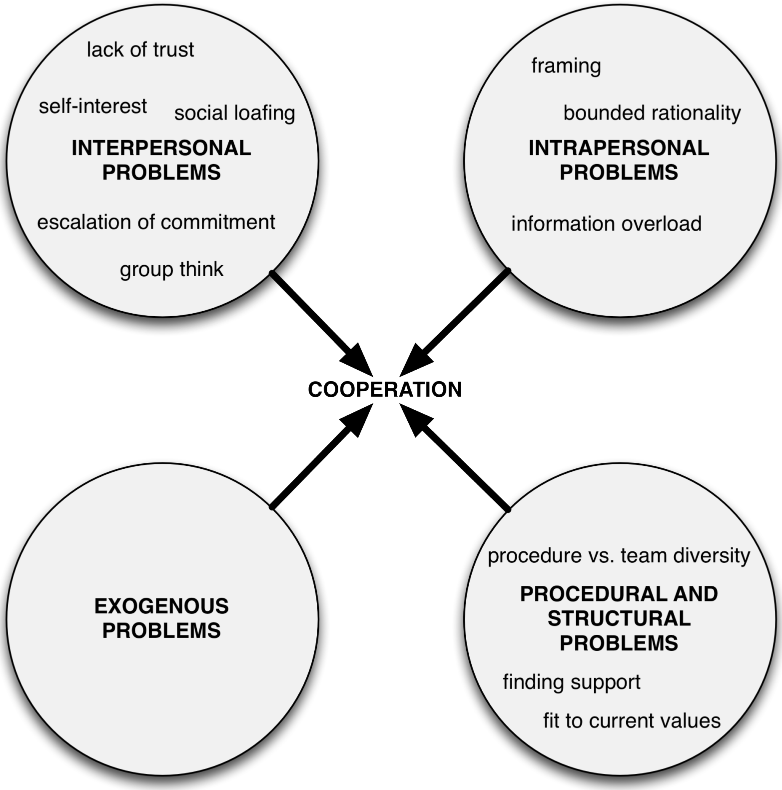


Figure 1.1 Four main types of problems in cooperation networks.

1.3.1 Intrapersonal problems

Kahneman and Tversky (1979) point out a *framing* effect when people choose to rather loose 4000 dollars with a probability of 80 percent than a 100 percent chance to loose 3000 dollars. In this case, they are risk seeking due to the negative way in which the problem is framed. A positively-framed problem - 80 percent chance of winning 4000 dollars or 100 percent chance of winning 3000 dollars - would result in risk averseness, because a sure win of 3000 dollars is preferred. LeBoeuf and Shafir (2003) elaborate on the framing effect by finding that deeper thought (longer thinking time) may decrease error in the decision making process.

There are numerous factors that we human decision makers have to take into account and, unlike computers, we cannot put them into a complex function that immediately tells us which person is best to cooperate with. Herbert Simon (Simon, 1982, 1991) came up with a notion called *bounded rationality* to denote such a phenomenon. We do not possess the cognitive ability to take into account each and every factor and solve the equation. Another phenomenon which is likely to occur as a result of such bounded rationality, is *satisficing*, a merger of the words satisfying and sufficing (Simon, 1982; Winter, 2000). What it means is that people come up with a solution that is good enough, but inherently non-optimal.

Another issue that a human decision maker faces is that a typical social network grows over time. As your network grows, the number of people that you can connect to increases, directly or indirectly. Typically, people have hundreds, or even thousands of people that they are connected to. If we count offline connections, we may even have more of them. Each of these contacts also has certain characteristics, or activities that they perform. Keeping track of them is practically undoable. In other words, we face an *information overload* (De Choudhury, Sundaram, John, & Seligmann, 2008). More specifically, based on the neocortical size of the human brain, Dunbar predicted that humans could only handle 150 persons in their social network (Dunbar, 1993). Based on empirical work, that number was adjusted to a mean social network size of 125 (Hill & Dunbar, 2003). To clarify, this means that in our daily lives, we on the average interact with some 125 people. So, if we meet person 126, we drop one among the now 126 from our social network, because cognitively seen we can only manage a social network of size 125.

1.3.2 Interpersonal problems

By nature, humans are self-interested (although not necessarily only so)(Whitworth & Whitworth, 2010). However, they always seek reasons for why they should cooperate (Crano & Prislin, 1995). Crano (1995) emphasises that vested interest is relevant here. When an individual personally feels the consequence, then the individual is more likely to show commitment. Colman and Pulford (2012) take a game-theoretic perspective to understand why people do or do not cooperate. In games with a definitive end, such as *one-shot games*, people tend to defect, whereas in games with no definitive end people tend to cooperate (Aumann, 1959).

Kogut (1989) found that joint ventures that have multiple relationships tend to be more stable. The main reason for this is *reciprocity*. The firms employ a so-called tit-for-tat strategy (Axelrod, 1984) in which they reciprocally reward technology transfer behaviour, and penalize competitive behaviour. Reciprocity in a learning network (Aviv & Ravid, 2005) occurs if a bidirectional link between persons A and B exists; person A communicates with B, and person B communicates with A. Nowak and Sigmund (2005) make a distinction between direct reciprocity (A helps B and B helps A) and indirect reciprocity (A helps B, B helps C and C helps A). They show that gossip may foster a good reputation and thus acts as indirect reciprocity.

Inter-firm cooperation often fails due to *free riding* behaviour. Increased group sizes and decreased cohesiveness are associated with increased free riding behaviour (Rokkan & Buvik, 2003; Liden, Wayne, Jaworski, & Bennett, 2004), also known as *social loafing* (Latané, Williams, & Harkins, 1979; Karau & Williams, 1993; Liden et al, 2004). Moreover, Chidambaram and Tung (2005) report that in computer-supported collaborative work, small groups outperform larger groups as a result of social loafing in larger groups.

Individual group members may face social pressure toward unanimity and loyalty to the group. Consequently, the group fails to weigh the risks and alternatives carefully, resulting in sub-optimal problems solving. This is also known as *group think* (Janis, 1982; Rose, 2011). The flip side of the coin shows that group members that have opposite preferences may take more radical decisions than the initial preferences showed. Such *group polarisation* (Moscovici and Zavalloni, 1969; Isenberg, 1986) is caused by social comparison or persuasive argumentation (Burnstein & Vinokur, 1977). People behave in a socially desirable way, but exaggerate in moving their point of view towards other members of the group (social comparison). Persuasive argumentation is the phenomenon that people exaggerate argument-finding for opposing perspectives, leading to polarisation of perspectives.

Escalation of commitment (Shubik, 1971; Ruthledge, 2011) occurs when people commit to their earlier action even though they have new information available that tells them their action is not optimal anymore. Groups tend to escalate commitment when they are held responsible for earlier time or money investments that were made (Ruthledge, 2011).

Lack of *trust* is an important threat to cooperation in networks. Trust is the expectance of cooperative behaviour of opponents, even when they do not meet again (La Porta, Lopez-de-Silanes, Shleifer, & Vishny, 1997). It is associated with performance of the government and large organisations (La Porta et al., 1997) and virtual teams (Rusman, Van Bruggen, Sloep, & Koper, 2009). When parties do not trust one another, they are likely to defect. In the Prisoner's Dilemma, in which two prisoners have the choice to cooperate or defect, they tend to defect because of a lack of trust, or reciprocity. Especially when they do not meet again – a *one-shot* *game* – defection is the option with the highest payoff. Dall'asta, Marsili and Pin (2012) argue that this very mechanism plays a role in cooperation networks.

1.3.3 Procedural and structural problems

When engaging in networked cooperation, people encounter various procedural challenges. For instance, the innovative process is very much dependent on fluent cooperation. The creative process can be described in several ways. Margaret Boden (2004) describes it as the exploration and transformation of existing ideas. Wallas (1976) distinguishes four stages of the creative process: preparation, incubation, illumination and the verification and expression of ideas. Osborn (1954) differentiates six stages: mess-finding (look for high level objective and goals), data-finding, problem- finding, idea-finding (divergent thinking), solution-finding (convergent thinking) and acceptance-finding. Schmid (1996) distinguishes four stages as part of the IPC-model: problem recognition, preparation, incubation and verification/elaboration. Each of these stages that these researchers describe have their specific challenges that we need to overcome.

The study by Bacharach (2005) raises the question when you should pursue diversity in a team, and when you should not. Depending on the question or work task at hand, we choose a more or less diverse team. For instance, coming up with novel solutions often requires a certain amount of creativity from a team. You may need different viewpoints, knowledge and skills to arrive at a novel solution. A team of diverse individuals may work in the creative process' divergent stage (idea generation), but the convergent stage (idea acceptance and implementation) may require more *homophily* (Ibarra, 1992) to achieve a common stance. Thus, it is important that a balance in diversity be kept, and roles in the team be fulfilled by the right individuals. One such role is the leader role; weak project leaders may be counterproductive for the success of the project (Pinto & Kharbanda, 1996). If strong leadership is absent, projects tend to become aimless and lose track, and meetings become indecisive.

In research and innovation implementation, it is important that you find the necessary support for the acceptance of your idea (Sie, Bitter-Rijpkema, & Sloep, 2010b). Reviewers of conference papers and journal articles and management of innovative firms should be aware of the value of your idea. One way of getting your idea accepted is borrowed from organisational change; a *guiding coalition* (Kotter, 1996) needs to be formed that supports the idea and that can persuade others. For example, the adoption of the Post-it was achieved by Arthur Fry, who gave the post-its to secretaries that adopted the Post-its and kept asking for more, even when his 'experiment' was over. Eventually, management was persuaded to take the Post-it into production. Also, a novel idea should fit the values of the stakeholders (Klein & Sorra, 1996).

1.3.4 Exogenous problems

Dignum (2002) stresses that cooperation, coordination and sharing in organisations "must be encouraged and nurtured". The need for a cooperative culture is emphasised by Shim and Steers (2012), who report that employees at Hyundai and Toyota consider a "'we' culture" to be key for cooperation, and consequently, organisational success.

Given the current economic crisis (2012), the Dutch government has decided to cut the budgets that they assign to the pubic libraries of the Netherlands. As a result, it was unsure whether the innovative learning network that we set up for the Dutch librarians, *Biebkracht*, could continue. Several studies report on the importance of funding for cooperation and innovation. For instance, funding plays an important role in research performance (Gulbrandsen & Smeby, 2005). Conversely, variations in funding schemes tend to have no effect on research performance (Auranen & Nieminen, 2010). Hanak and Rueben (2006) draw attention to the importance of funding for innovation in transport.

1.4 Main research questions

The above discussion inventories a host of stumbling blocks for cooperation to get off the ground. Actually, the number of problems is too large for one thesis to tackle. In this thesis, we will limit ourselves to a subset of problems that may all be subsumed under the following main research question:

How can we assemble individuals that want to cooperate to create something new?

This main question has a number of aspects, each associated with a question. A team of experts assigned to solve a particular problem should reflect all types of knowledge that is needed to do so. Furthermore, the team needs to be able to work together. That is, their behaviours should be compatible. There are various factors that influence cooperation in networks in positive and negative ways, and they should play together nicely. We define question 1 as follows:

1. What factors influence cooperation between individuals?

It is important that we take into account both perspectives of individuals involved in cooperation: the practitioners, and the experts. From a practitioner's perspective, we elicited knowledge in personal, professional learning networks (Chapter 2), and from an expert perspective, we focused on elicitation of knowledge about cooperation in networks (Chapter 3). The two contexts served as a triangulation of the literature review that was performed at the start of this thesis' study. They lead to the following two subquestions:

1a. What factors do practitioners perceive to influence cooperation between individuals?

1b. What factors do experts perceive to influence cooperation between individuals?

Also, one needs to know about the interplay between these factors; how they influence one another. In Chapters 4 and 5 we present computer models that simulate the behaviour of the factors that influence cooperation. We study how the factors interact with each other and how their interaction changes when varying social network size and network density (Chapter 4). Also, we study more elaborately how sensitive the model is to changes in the factors (Chapter 5). That is, for each factor, we vary its value within a predefined range and measure it repeatedly during simulation, yielding 1450 simulation runs. The following subquestion to question 1 is investigated in Chapters 4 and 5:

1c. How do the factors that influence cooperation interact with one another?

During the creative part of the process, you typically need diverse views from individuals, to create that new perspective that is needed to create something new, or innovative, or appealing. Though, innovation does not merely consist of being creative. It also involves implementation of your new product (Denning, 2012). That is, unless you are given a bag of money unconditionally, you need to persuade others of the value of your idea or product. Consumers need to buy and use your product. This raises the following question:

2. How can we persuade individuals to cooperate so that their ideas will be accepted and implemented?

Sometimes you need to persuade others in advance to actually receive the money to work on a product, sometimes you need to show others your new product and try to persuade them afterwards. While in the innovative process, you might want to involve that mad scientist that can do exceptional things, but at the same time is unable to communicate his ideas to management. Creative individuals are not always the right people to persuade others. Thus, you need someone that has the ability to persuade others, or someone that has enough power to force decisions. Also, someone that has a certain reputation or status could be welcome in your team, as this eases the acceptation and adoption of your product or idea. The above leads to the following subquestion to question 2:

2a. How do we define someone having the ability to persuade others?

One of the aims of the work performed in this thesis was to support the innovative process by means of a system that brings together individuals. Such a system should base a recommendation of future partners or alliances on the knowledgeability and persuasion skills of peers. We were interested in how users perceive the functioning of the system, that is, is it able to recommend peers that can boost the implementation or acceptance of an idea? In particular, two of our studies focused on recommendation of peers that can help implement a research idea. We extracted and analysed a co-authorship network in order to recommend future co-authors. It is sometimes the case that these recommended co-authors are not known to the user. How do they cope with this? How do they perceive such a recommended co-author? In Chapter 6, we present a first version of the COCOON system that recommends future co-authors. It addresses the following subquestion:

2b. What co-authors do users prefer to be recommended: just the people that they have already worked with, or also new co-authors?

Naturally, we also want to study what the value of a recommendation of co-authors itself is. In Chapter 7, we present a second version of the COCOON system, called CORE (CO-author REcommendation). CORE aims at finding both influential peers and knowledgeable peers to foster implementation of a research idea. Users can choose themselves how they balance between influential and knowledgeable peers. Chapter 7 addresses the following and final subquestion:

2c. How do users value recommendations of future co-authors based on their influence and like-mindedness?

To clarify the above questions, in particular their interplay, Figure 1.2 lays out the structure of this thesis.

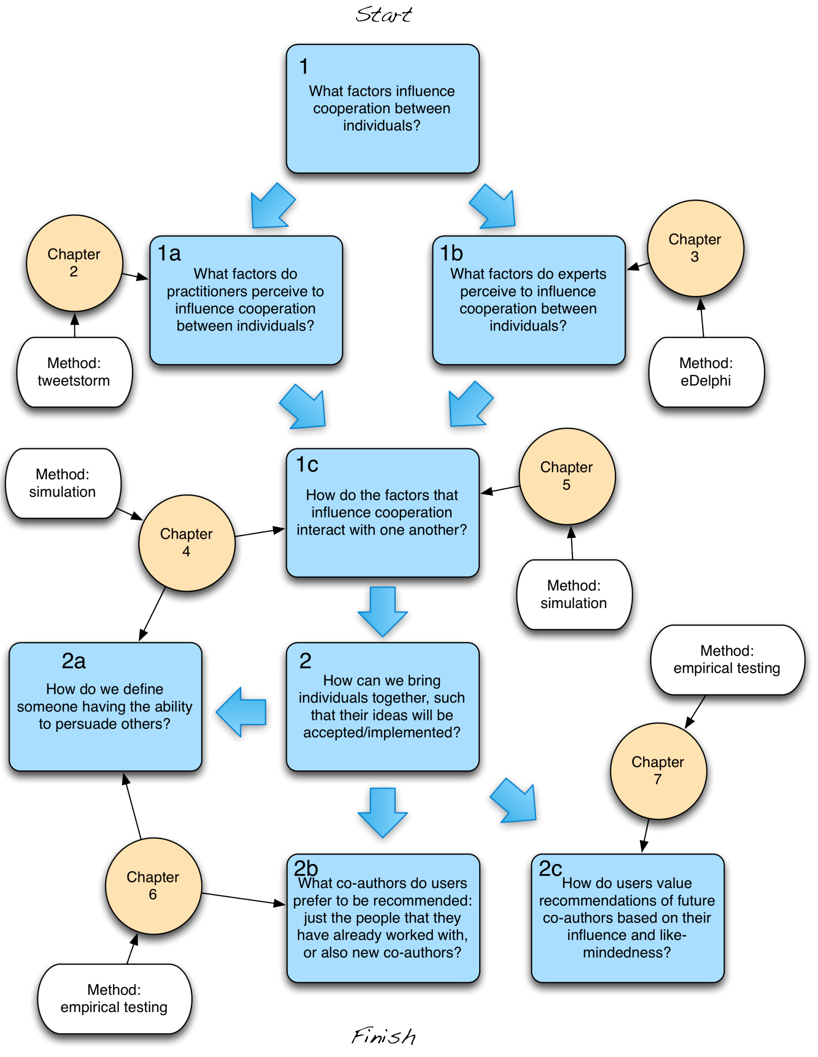


Figure 1.2. Overview of this thesis' structure.
Blue rectangles represent the research questions posed in this chapter, which are discussed in the subsequent chapters (beige circles); each chapter employs a specific research method (white rounded rectangles).

CHAPTER 2

Goals, Motivation for, and Outcomes of Personal Learning through Networks: Results of a Tweetstorm

In order to offer help in cooperation networks, we first need to know what constitutes a cooperation network. We need to know, for example, how individuals interact, how they cooperate, what they value in cooperation. This chapter investigates how practitioners perceive their engagement in cooperation networks by studying a particular kind of cooperation networks: personal, professional learning networks.

We asked a group of professional learners to provide us with the contacts that they learn from in their daily professional lives. We also asked them how they connected to their contacts; through social media, email, or face-to-face. Afterwards, we employed a novel type of knowledge elicitation, the Tweetstorm, which is a merger of Twitter and the brainstorm technique. 'Tweets', messages constrained by a 140-character limit, are perfectly suited to generate short statements (brainstorm) about how they perceive their involvement in a learning network, and how they gain value from it.

This chapter is based on: Sie, R.L.L., Pataraia, N., Boursinou, E., Rajagopal, K., Falconer, I., Margaryan, A., Bitter-Rijpkema, M., Littlejohn, A., Sloep, P.B. (submitted). Goals, Motivation for, and Outcomes of Learning through Networks: Results of a Tweetstorm.

Abstract

Learning networks are no longer designed just by moderators. Recent developments in the use of social media for learning have put the learner in the driver's seat. Learners consider their goals, motivations and expected outcomes before designing their personal learning network. Previous research focused on the factors that influence learning in electronic environments, but these studies were mainly conducted in an era in which online social media were not yet used to design personal learning networks. The current paper reports findings of a study that examined factors impacting professional learning through networks. A personal learning network identification session and a brainstorm via Twitter (*Tweetstorm*) regarding goals, motivational factors and outcomes of learning through networks were conducted. Based on the analysis, the article concludes that seven factors play a pivotal role in personal, professional learning through networks: sharing, motivation, perceived value of the network, feedback, personal learning, trust and support, and peer characteristics and peer value. Also, in motivation, different perspectives, motivation, social media and collaboration, reciprocity, intrinsic motivation, innovation, status and reputation and networking strategies play an important role. Future work focuses on investigating the interplay between factors that influence networked learning that are identified in this article.

2.1 Introduction

Social capital theory states that "valued resources and expertise are embedded within social networks" (Penuel, Riel, Krause, & Frank, 2009, p.126). Networks serve multiple purposes and different types of network relationships lead to different network outcomes (Finkelstein & Lacelle-Peterson, 1992; Pifer, 2010). For instance, social networks can act as communication channels through which knowledge is disseminated (Rogers, 1995; Owen-Smith & Powell, 2004). However, networks are perceived not only as channels for the transfer of knowledge but also as vehicles for the creation of new knowledge through a process of collective sense making (Ring & Van de Ven, 1994). Various types of connections and flows link network members to one another, such as information, materials, resources, services and social support (Borgatti & Cross, 2003).

In recent years, research findings have documented the importance of a network perspective for learning (Sie et al., 2012; Dawson, Bakharia, & Heathcote, 2010; Haythornthwaite & De Laat, 2010; Berlanga, Bitter-Rijpkema, Brouns, & Sloep, 2008a; Siemens, 2006; Sloep & Berlanga, 2011). The social interactions that take place during learning constitute a *learning network* (Downes, 2010; Sloep, Van der Klink, Brouns, Van Bruggen, & Didderen, 2011). In a learning network, learners are represented as nodes, and their learning interactions are represented as the edges between the nodes. Paths in the network may be regarded as a relationship between learners. Also, the term 'learning network' is often used to refer to the data extracted from interactions in online collaboration environments, such as Personal Learning Environments (PLEs). PLEs are a new set of technologies, mainly social media, meant to guide the assessment and recognition of learning (Attwell, 2007). Also, PLEs aim to assist learners in sharing and merging content from several sources (Ebner, Schön, Taraghi, & Drachsler, 2011).

If the reader considers the individual learner's personal preferences and characteristics with a view to generate learner-specific content and connections, it is called a *personal learning network* (PLN). Yet, very little is known about what exactly characterises learning in a PLN. Especially in an era in which social media are gaining popularity as a means of learning (e.g. Ebner et al., 2011), it is important that one investigates how people learn, and how they create a balance between the use of offline contact and online social tools. Väljataga and Fiedler (2009) emphasise that learners should be able to adapt their use of social media to particular learning activities. To assist such learners, we need to know what constitutes a learning tie (Haythornthwaite & De Laat, 2010). One needs to know whom people learn from, what they learn, how they learn and what drives them to learn. Specifically, one needs to know what tools learners use while engaging in learning networks and one needs to explore their 'networking attitude' (Rajagopal, Joosten-ten Brinke, Van Bruggen, & Sloep, 2012).

2.1.1 Related work

The question *whom we learn from* has a long history in educational research and several learning theories aim to capture the social process of learning. Bandura (1977) defines social learning as learning from others; modelling and imitating others' behaviour. Vygotsky (1978) underlines that learning, internalising behaviour, occurs by imitation; we learn from others by example. Wenger (1998) contends that learning is practice-driven; people share a common interest or practice. Learners influence and learn from one another as they engage in their "community of practice". Connectivism (Siemens, 2005), a theory that explicitly refers to learning with technology, claims that "learning is a process of connecting to specialized nodes or information resources". This includes learning from objects, or organizations that possess knowledge.

Dillenbourg (1999, p.2) defines that we learn collaboratively by having "a situation in which two or more people learn or attempt to learn something together". Four main types of activities are distinguished to describe *how* we learn at the workplace (Eraut, 2004): 1) participation in group activities, 2) working alongside others, 3) tackling challenging tasks, and 4) working with clients. The first, second and fourth point towards social, collaborative actions, which may be important for our understanding of personal, professional learning networks.

What we learn in the workplace ranges from task performance, awareness and understanding, personal development, teamwork, role performance, academic knowledge and skills, decision making and problem solving, and judgement (Eraut, 2004). Roger Schank (1995) states that we internalise so-called *scripts* of consecutive actions when we learn by doing. This is similar to the social learning view of Bandura (1977), who claims that we learn from others by constructing a model of what others do and try to imitate this.

The reason *why* learners engage in learning networks may be that they share a common interest or practice (Lave, 1991), are keen to exchange of ideas (Pirolli, 2009) and want to receive and provide support (Fetter, Berlanga, & Sloep, 2010; Berlanga, Sloep, Kester, Brouns, Van Rosmalen, & Koper, 2008b; Van Rosmalen et al., 2007). They also call on each other when they have a problem to solve or knowledgeability to offer (Dekker & Kingma, 1999). Social support theories posit that network relationships offer both instrumental and emotional support to network members (Gerstick, Bartunek & Dutton, 2000). Instrumental relationships encompass resources such as professional advice, information, and expertise, whereas emotional relationships provide encouragement, friendship, support and ways of communicating information (Ibarra, 1993). Access to knowledge resources may guide learner engagement in learning networks (Hollingshead, Fulk, & Monge, 2002). Also, learner engagement is subject to the learner's interest (Billett, 2004).

2.1.2 Outline

Ibarra, Kilduff & Tsai (2005) underline that much has to be learnt about how people use, adapt and change their networks of relationships. We conducted a study to investigate what characterises learning in a personal learning network. We focussed on professional learners in particular, as they are likely to constitute the majority of PLN users (Sloep et al., 2011). This resulted in the following research question:

How do learners construct, use and perceive their personal, professional learning networks?

The study attempts to increase our understanding of how moderators and learners design professional, personal learning networks; it does so by exploring how professionals utilise their networks. We present findings from a new type of knowledge elicitation, the *Tweetstorm*. The Tweetstorm is an online, open brainstorm session via Twitter, a microblogging platform. In advance of the Tweetstorm session, we charted the egocentric networks – the network as seen from the perspective of an individual - from a group of researchers interested in personal learning environments (PLN identification session), to provide a context.

The present chapter starts off with the way we collected data and how we went about conducting the experiment for both the PLN identification session and the Tweetstorm. Subsequently, we present and discuss the results of the PLN identification session and the Tweetstorm. Finally, we will outline some conclusions and provide some suggestions for future work.

2.2 Method

2.2.1 Participants

2.2.1.1 PLN identification session

Participants were chiefly educational researchers with an interest in Personal Learning Environments. Typically, a conference allows researchers to publicise themselves, but also to maintain and expand their existing network. More importantly, researchers learn from each other during a conference. The latter relates directly to our aim: to identify the contacts that professional learners in a network learn from, and the goals and motivation for their social learning behaviour.

A total of six participants (active in educational research) took part in the PLN identification session, which was part of a workshop at the PLE conference. The workshop was announced before the start of the conference. Their main characteristics are provided in Table 2.1. No inducement was offered for their participation.

Table 2.1. Overview of the participants' main characteristics

ID	gender	age range	profession	discipline
1	m	35-44	PhD student	education
2	f	45-54	teacher	cultural and ethnic studies
3	m	35-44	professor	other
4	m	25-34	post-doc	education
5	m	25-34	PhD student	computer sciences
6	f	25-34	teacher	sociology

2.2.1.2 Tweetstorm

Due to the public nature of Twitter, the Tweetstorm was open to anyone who was interested and managed to spot it. A total of 31 participants actively engaged in it by *tweeting* (uttering statements called 'tweets') or *retweeting* (forwarding tweets). These included the six participants that participated in the antecedent PLN identification session. The Tweetstorm was announced through the website of the

PLE conference. The use of Twitter meant that we could only identify participants by their Twitter username (quasi-anonymity). As indicated, passive, read-only participants ('lurkers') could also join the Tweetstorm. As Twitter does not allow for tracking of 'reads', lurkers could have (indirectly) influenced the Tweetstorm by discussing with active participants offline. No inducement was offered for participants' cooperation.

2.2.1.3 Statement sorting

We invited a group of experts to participate in a sorting experiment to independently categorise the statements that were extracted from the tweets. Since the statements were about learning in networks, 34 experts from affiliated universities, researchers in the educational domain, were invited via email, of which nine responded positively (seven females, two males). Their occupation varied from PhD student to associate professor. Again, no inducement was offered for their help.

2.2.2 Materials

2.2.2.1 PLN identification session

A custom-built online environment (PLN identification tool) was used in which participants could register themselves and identify the contacts in their PLN (Figure 2.1). The PLN identification session lasted 45 minutes in total. The environment was accessible through the Internet URL 145.20.132.20/rse/test/page/PLE. For ease of use, the URL given to the participants was shortened using an online service called Bit.ly. The environment was tested during a pilot session at Glasgow Caledonian University. Five participants, all educational researchers, tested the environment and were given the opportunity to 1) reflect on clarity and usefulness of the questions, and 2) to provide suggestions for improvement. As a result, the survey instruments and questions were refined prior to the actual session. Although some of the answer options that were added seem to overlap with the existing ones, the test participants felt these needed to be added. For instance, 'external colleague' and 'research collaborator' may have overlap in meaning.

Participants could edit or delete the contacts that they entered (bottom of Figure 2.1; actual entries are left out for privacy reasons).

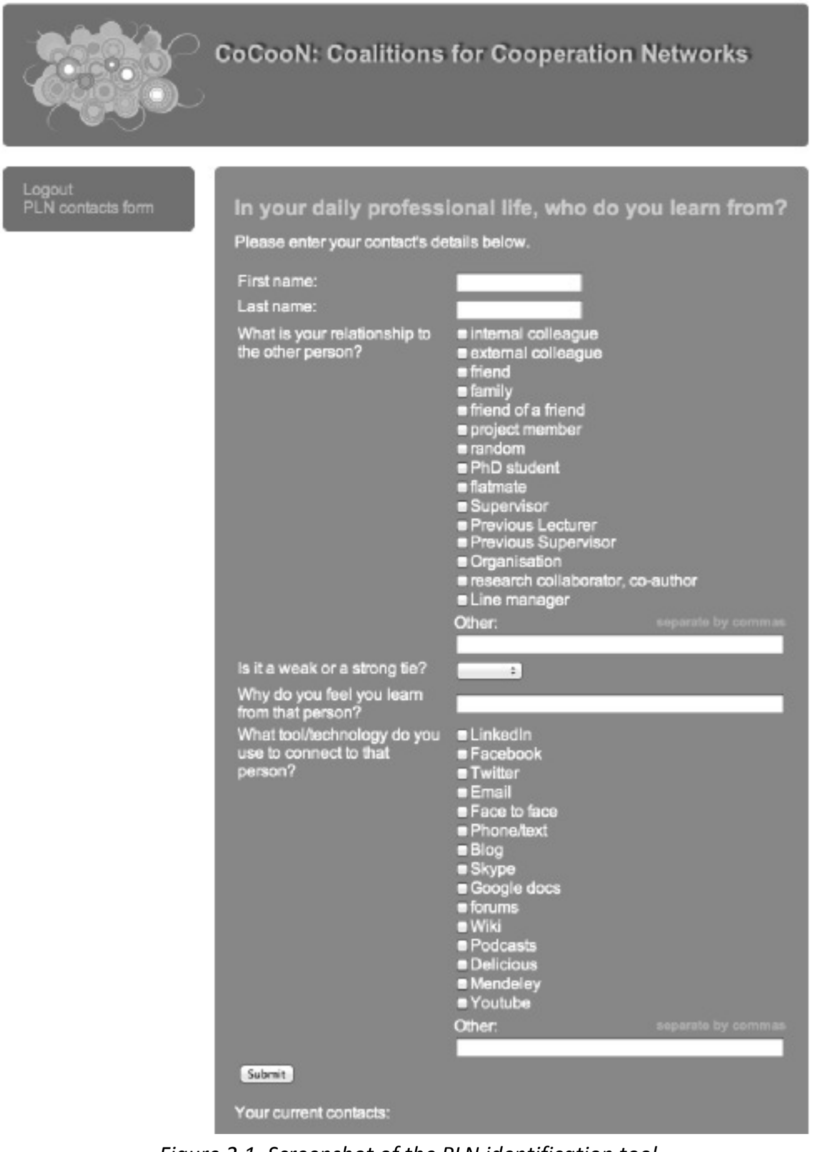


Figure 2.1. Screenshot of the PLN identification tool.

The PLN identification session was analysed in SPSS. The Tweetstorm was analysed using the card sorting tool Websort.net (http://www.websort.net); cluster analysis was performed using the multidendrograms software package (Fernández & Gómez, 2008).

2.2.2.2 Tweetstorm

A custom-created hashtag #plntweet and a twitter account @PLNtweetstorm were created in advance of the workshop to guide the process, in order to post trigger questions.

During the Tweetstorm, a so-called *twitterwall* was shown at the workshop venue. Such a twitterwall allows that an overview of all tweets with the same hashtag, in this case #plntweet, be presented to all participants. Besides, the twitterwall allowed for easy aggregation of the tweets for analysis. Figure 2.2 shows a part of the #plntweet archive in Twapperkeeper twitterwall (http://www.twapperkeeper.com).

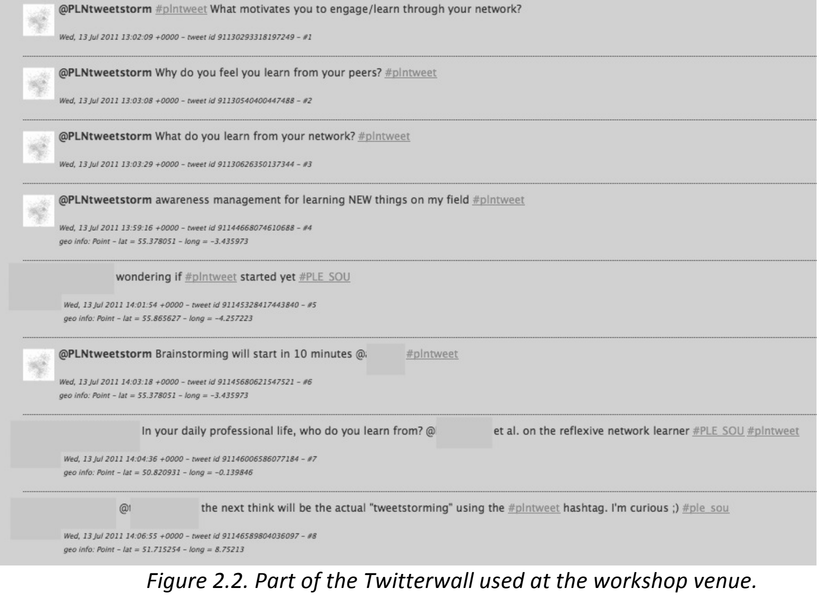


Figure 2.2. Part of the Twitterwall used at the workshop venue.

2.2.2.3 Statement sorting

The statements that resulted from the tweets were categorised by expert educational researchers using a tool called Websort.net, which is designed to do card sorting experiments and corresponding data analysis. Having the statements in digital form allows for card sorting online. The main advantages of online card sorting systems are: 1) there is no need to organise a face-to-face expert session, 2) experts can sort statements anonymously, 3) experts can participate at distant locations and 4) fast data aggregation and analysis. WebSort provides a number of data aggregation (e.g. items vs. items, items vs. categories) and visualisation methods (e.g. tree structure, tables). Participants are not able to see each other's categorisations. Also, the categorisation did not have any time-constraints.

The multidendrograms software package (Fernández & Gómez, 2008) was used to perform agglomerative hierarchical cluster analysis (AHCA) with complete linkage

(Defays, 1977) to find core clusters of statements. AHCA starts with all statements in distinct clusters. In subsequent iterations, clusters are merged based on their similarity, until the appropriate number of clusters is reached. That is, the resulting clusters should be roughly equal in diameter, the maximum distance between two items in a cluster. Merging takes place if the average distance between two clusters is small (complete linkage). In the beginning, cluster distances are inherently small, as every statement has its own cluster. The similarity of statements is based on the number of times two statements co-occur in categories defined by the experts. For instance, if expert 1 puts statement A and B in a single category and expert 2 puts statement A and B in a single category, then the similarity between statement A and B increases. Similarity calculation is category-name-independent. Consequently, if all experts put statement A and B in the same category, but name the category differently, similarity is still 100%.

2.2.3 Procedure

The experiment was conducted at the Personal Learning Environments conference (PLE 2011) in Southampton (http://www.pleconf.com), during a workshop. We employed a two-phase approach to collect data. First, to provide a clear context in advance, we offered participants the opportunity to reflect on and articulate their own learning networks by naming at least ten people or organisations they learn from in their daily professional life (PLN identification session). Second, a Tweetstorm session was held, in which participants were asked to use their Twitter accounts to contribute to the discussion.

2.2.3.1 PLN identification session

At registration, participants of the PLN identification session described their profile in terms of their age range, gender, occupation, discipline and work experience. The main advantage of providing and keeping login credentials is that participants can be asked to identify contacts at a later point in time (repeated measure), to see how their network and perception of this network evolves.

After registration, participants could add contacts that they learn from through the PLN contacts form. For each contact, the participants had to answer the following questions:

1. What is your relationship to the other person?
2. Is it a weak or a strong tie?
3. Why do you feel you learn from that person?
4. What tool/technology do you use to connect to that person?

Although participants were asked to identify their learning contacts, the relationships between contacts and contacts' characteristics were not analysed. Using SPSS statistical software version 18, we calculated averages per type of contact and tool that learners used to connect to their learning contacts.

2.2.3.2 Tweetstorm
The moderators (three) tried to trigger participants by posting three main questions about PLNs to Twitter using the #plntweet hashtag:

1. What motivates you to engage/learn through your network?
2. Why do you feel you learn from your peers?
3. What do you learn from your network?

Participants were asked to add the hashtag #plntweet to each and every one of their tweets to make sure the results could be aggregated after the Tweetstorm had ended. The Tweetstorm lasted 45 minutes in total.

2.2.3.3 Statement sorting
The tweets were aggregated and split up into smaller pieces of information, as most of the tweets addressed multiple questions at once. That is, one tweet could answer both the question what motivates the learner and what the learner learns through the network. As the researchers posted (tweeted) the triggering questions separately, it was not expected that participants would answer multiple questions in a single tweet. Therefore, tweets were split up into statements that answered a single triggering question. Moreover, some of the answers contained distinct parts that could possibly be interpreted and categorised differently from each other. For example, one part of the answer could be about feedback, whereas another part could be about inspiration. After splitting up these tweets into separate statements, we uploaded these in the Websort.net environment. Following this, we asked the experts to categorise the statements. To prevent researcher bias, no pre-defined categories were provided. Experts could define and name categories themselves.

We used the Websort environment to export the sorting data to two types of results. First, we exported the summary for the categories that the experts identified. Second, since little overlap was found (inherent to the fact that experts could name the categories themselves), we needed to analyse the overlap using agglomerative hierarchical cluster analysis. Therefore, we exported the data to an *item-item* similarity matrix. This matrix is too large to be reported here in full, however it is available on http://www.open.ou.nl/rse/Rory_Sie/Downloads.html. Finally, AHCA with complete linkage was performed to find core clusters of statements.

2.3 Results

2.3.1 PLN identification session

2.3.1.1 Whom do participants learn from?
Fifteen types of connections and fifteen different tools for communication were identified in the answers by the participants of the introductory session (Figure

2.3). From the six participants, one participant had named only five contacts. The rest had identified more than ten contacts, ranging from ten to twenty-four. In total, 261 contacts were identified. The participants could be connected to the same peer by more than one type of connection or tool. For example, a research collaborator could also be the participant's friend and use face-to-face as well as email communication.

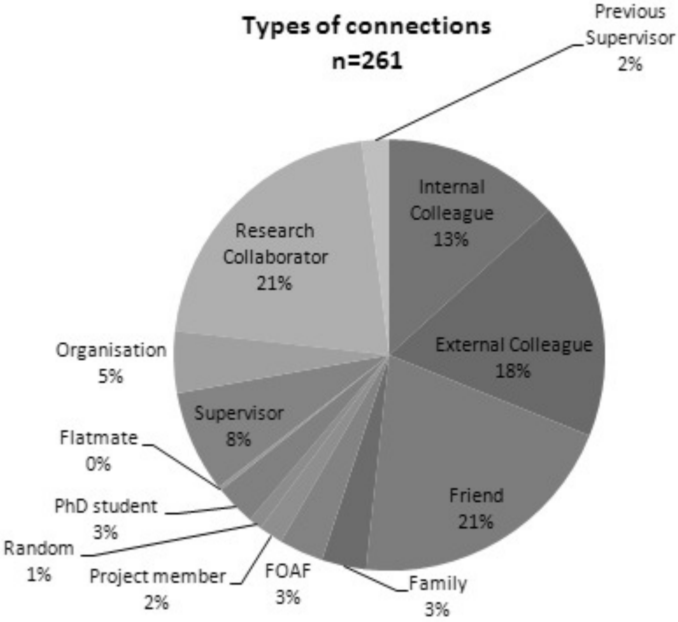


Figure 2.3. Whom do people learn from?

The findings revealed that the most common type of relationship in a learning network was research collaborator, friend and external colleague. 40% of research collaborators were at the same time friends. Following in order of meaningful connections were internal colleagues and supervisors.

2.3.1.2 What tools do they use?

In total, thirteen out of fifteen distinct tools were selected by participants (Figure 2.4). The tools used most commonly were Twitter (18%, per participant: M=.68, SD=.47), email (19%, per participant: M=.65, SD=.48) and face-to-face communication (18%, per participant: M=.65, SD=.48). Although the social bookmarking tools Delicious and Wikis were an option, they were never mentioned.

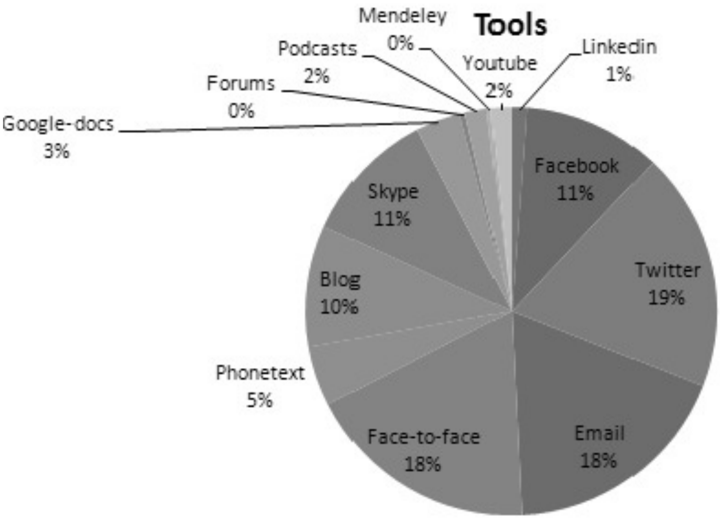


Figure 2.4. Tools used to learn from peers.

2.3.2 Tweetstorm

Participants posted a total of 139 tweets (M = 4.48; SD = 6.28) (38 retweets) with the requested #plntweet hashtag. Sorting of the tweets entailed that we had to remove retweets, triggering questions, and split up tweets with multiple statements in them. A total of 83 statements were extracted from the Tweetstorm (see http://www.open.ou.nl/rse/Rory_Sie/Downloads.html).

2.3.3 Statement sorting

There was no time-constraint set for the sorting exercise. Experts spent 51 minutes on average sorting (SD = 35). Table 2.2 shows the categorisations by the experts.

Table 2.2. Categorisation by experts.

Category	**Experts**	**Total items**	**Unique items**	**Agreement**
(Learning) benefits	1	24	24	1
Advantages	1	16	16	1
And take	1	9	9	1
Autonomy	1	1	1	1
Balance between give and take. Economic/rational approach	1	3	3	1
Based on a negative attitude	1	2	2	1
Characteristics of PLN	1	13	13	1
Characteristics/features of a network	1	12	12	1
Collaboration and community	1	5	5	1
Collaborative learning (with peers)	1	8	8	1
community identity, less relevant for me	1	4	4	1
Competences needed to be part of a	1	4	4	1

Network				
creation of a community of learners	1	13	13	1
definition of a network	1	16	16	1
Different conceptions of a PNL	1	15	15	1
Difficulties/problems	1	2	2	1
diversity	1	3	3	1
don't agree	1	5	5	1
effectiveness	1	3	3	1
efficiency	2	2	1	1
Expectatives	1	11	11	1
experiences	1	6	6	1
Feedback	1	4	4	1
Fun, happiness	1	3	3	1
fun, passion	1	13	13	1
General benefits of learning in a network: acquiring reputation/status based on quality of ideas	1	6	6	1
General benefits of learning in a network: efficiency/easiness/efficacy	1	2	2	1
General benefits of learning in a network: motivation/inspiration/passion	1	5	5	1
General benefits of learning in a network: quality/diversity/newness of ideas/perspectives	1	5	5	1
General benefits of learning in a network: rolemodeling/examples/(common)reference framework	1	11	11	1
General benefits of learning in a network: supporting each other	1	6	6	1
General benefits of learning in a network: tailored to personal learning needs	1	4	4	1
Getting the world inside	1	12	12	1
Getting your world outside	1	5	5	1
Give	1	6	6	1
Goals	1	2	2	1
hmmm	1	1	1	1
hype	1	7	7	1
I don't understand :(	1	2	2	1
Ideas, information, inspiration and opinions	1	19	19	1
innovation	1	1	1	1
instruction	1	8	8	1
Interaction and support	1	6	6	1
interpretations	1	1	1	1
intrinsic motivation	1	8	8	1
intrinsic motivation from connecting to people	1	8	8	1
Knowledge, expertise	1	10	10	1
learning by interactions	1	23	23	1
learning goal	1	5	5	1

learning in networks	1	2	2	1
learning mainly as social learning=social exchange	1	13	13	1
learning to learn	1	9	9	1
learning=individual benefit receiving	1	39	39	1
limitations	1	1	1	1
maintain relations	1	3	3	1
make work interesting and inspirational	1	27	27	1
Misconceptions	1	5	5	1
models and expertise	1	6	6	1
Motivation	2	21	14	0.75
motivation: give and take	1	1	1	1
Motivations to be part of a Network	1	9	9	1
opinions	1	3	3	1
passion	1	2	2	1
pathetic statements	1	3	3	1
peers	1	3	3	1
People in My Network	1	13	13	1
perceived support by the network	1	12	12	1
Personal development	1	2	2	1
personal drive	1	7	7	1
personal gains by the network of learners	1	29	29	1
Personal learning due to participation in a network	1	12	12	1
platitudes	1	2	2	1
Problem solving and ask for help	1	6	6	1
Realtime interaction	1	3	3	1
Reasons for PLN	1	12	12	1
Reasons of learning (general)	1	3	3	1
Reflection and feedback often with peers	1	11	11	1
relying on others	1	14	14	1
reputation	2	6	5	0.6
resources	1	10	10	1
Roles	1	3	3	1
self-confidence	1	1	1	1
sharing	4	36	23	0.39
Social, informal interaction	1	5	5	1
Status	2	11	7	0.79
Stay in touch, connecting	1	5	5	1
Stay up-to-date	1	4	4	1
Support	1	3	3	1
trust, secure	1	3	3	1
Twitter	1	2	2	1
use network strategically	1	19	19	1
use of ICT	1	6	6	1

The column ‘Experts’ represents the number of experts that gave a category each particular name. For instance, ‘sharing’ was named as a category by four experts. The column ‘agreement’ shows to what extent the experts that named that category also put the same statements in that category. As Table 2.2 shows, nearly no overlap in category names was found. The reason for this is clear and expected; the experts could define the names for the categories themselves.

Figure 2.5 provides the results of the agglomerative hierarchical cluster analysis. The statements are coded, and can be found at http://www.open.ou.nl/rse/Rory_Sie/Downloads.html . The results can be interpreted in several ways, following the (agglomerative) nature of this method. For instance, on the lowest level seven clusters can be found (Appendix A). Cluster 1 was named *sharing* and included five statements. An example of such statements included "*sharing is key*". Cluster 2 was named *motivation* and included 32 statements. To exemplify, one statement mentioned "*Learning with others is more rewarding and rich than on your own*". Cluster 3 was named *Perceived value of the network* and included sixteen statements of which "*Finding out about latest research*" was one of them. Cluster 4 was named *feedback* and included four statements such as "*Feedback on thoughts and ideas*" and "*Instantaneous feedback, news, useful links, arguments and opinions*". Cluster 5 was named *personal learning* and comprised eleven statements. Cluster 5 included, for example, the statement "*Using my network to find information and learn is the most effective and fast way to get the things I need*". Cluster 6 was named *Trust and support* and comprised nine statements. Examples of these statements include "*Ask for help and they will engage and help me*" and "*I can also discuss some of the concerns and insecurities I have within a peer group informally*". Especially the latter emphasises the need for a trusted, informal support structure. Cluster 7 was named *peer characteristics and value* and included statements about how peers contribute to the participants' learning. Statements include "*Members of my PLN are very intelligent, inspirational, insightful and innovative*" and "*The people I learn from are passionate, critical and informed. They are my role models learners* [sic] *in this digital age*".

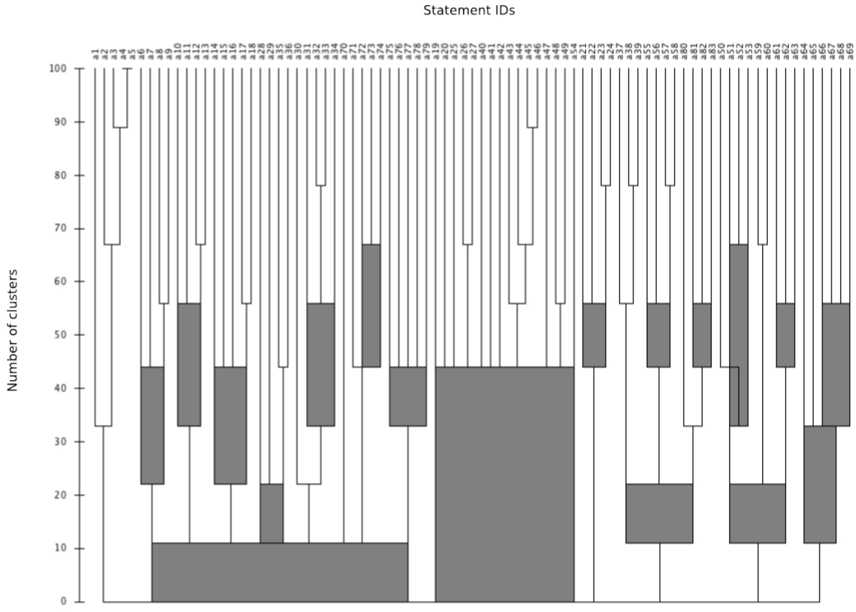


Figure 2.5. Results of hierarchical cluster analysis.

On the next level, fourteen clusters were found. The initial seven clusters remained the same, except for the cluster motivation, which could be split into eight subclusters (Table 2.3):

- *Different perspectives* (e.g. “Learn from your peers - "Views I hadn't considered, opinions I disagree with, ideas that inspire me"”),
- *Motivation* (e.g. “For me, learning through my network is the most fun way of learning”),
- *Social media and collaboration* (e.g. “Twitter is a fine balance between the personal and the social. No-one learns in a vacuum, but we all learn uniquely”),
- *Reciprocity* (e.g. “Conversation is 2-way. I can give to my network as well as take from it”),
- *Intrinsic motivation* (e.g. “I use my PLN because of the autonomy it provides me”),
- *Innovation* (e.g. “By results collaboratively achieved - new methods under construction e.g. by MOOC ing. Old scales don't work.”),
- *Status and reputation* (e.g. “Not everyone has equal status in my PLN”) and
- *Networking strategies* (e.g. “My PLN allows me to connect to new people, communities and artefacts”).

The other clusters remained the same, resulting in fourteen clusters in total. For clarification purposes, Figure 2.6 shows the seven core clusters and their subclusters.

Table 2.3. Statements per cluster at the level of fourteen core clusters.

cluster	name	statements
1.1	Sharing	a1, a2, a3, a4, a5
2.1	Different perspectives	a6, a7, a8, a9
2.2	Motivation	a10, a11, a12, a13
2.3	Social media and collaboration	a14, a15, a16, a17, a18
2.4	reciprocity	a28, a29, a35, a36
2.5	intrinsic motivation	a30, a31, a32, a33, a34
2.6	innovation	a70
2.7	status and reputation	a71, a72, a73, a74
2.8	networking strategies	a75, a76, a77, a78, a79
3.1	Perceived value of the network	a19, a20, a25, a26, a27, a40, a41, a42, a43, a44, a45, a46, a47, a48, a49, a54
4.1	Feedback	a21, a22, a23, a24
5.1	Personal learning	a37, a38, a39, a55, a56, a57, a58, a80, a81, a82, a83
6.1	Trust and support	a50, a51, a52, a53, a59, a60, a61, a62, a63
7.1	Peer characteristics and value	a64, a65, a66, a67, a68, a69

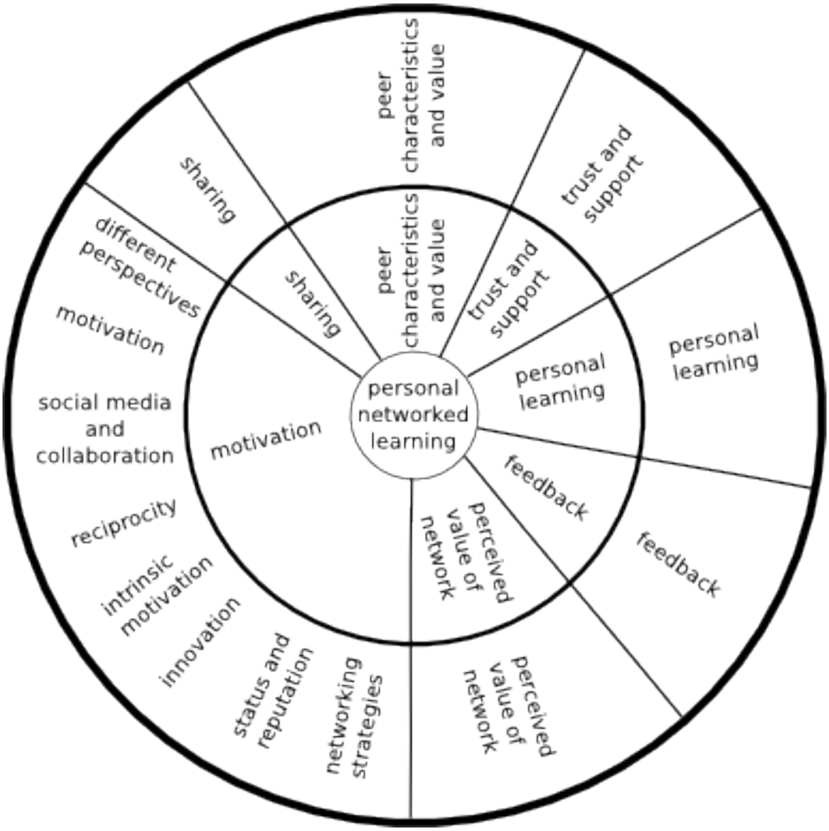


Figure 2.6. Seven core clusters and their fourteen subclusters.

2.4 Discussion

The PLN identification session, which focused on identification of egocentric networks, revealed some interesting findings. First, we found that the participants learn mainly from research collaborators, friends and external colleagues. For this, they used face-to-face, email and Twitter as main modes of communication. The Tweetstorm and the corresponding agglomerative hierarchical cluster analysis resulted in a core set of seven clusters and fourteen subclusters. At the level of the seven clusters, the cluster ‘sharing’ is consistent with research by Olson, Grudin and Horvitz (2004, p.1) who state “*Information sharing is of immense value in the workplace because it reduces duplication of effort, and sits at the foundations of collaboration*”. Also, Swan (2002) stresses the importance of interaction for teaching and learning in a network. On the other hand, Fogel and Nehmad (2009) report that the majority of men and women included a picture of themselves in their profile, but did not share their phone number and home address. Thus, people only share personal information to a limited extent. These two opposing views support that trust (cluster 6) is important in a personal learning network, but also calls for a balance between information sharing and trust. Furthermore, the importance of trust and support for learning is partly supported by Lankau and Scandura (2002), who contend that there exists a positive relationship between vocational support (mentoring in the workplace) and personal learning. In that same study, it was found that roles are an important indicator for skill development, which supports our findings that 'peer characteristics and value' play a part in personal learning networks.

Ames and Archer (1988, p.264) report that "*a mastery goal orientation may foster a way of thinking that is necessary to sustain student involvement in learning as well as increase the likelihood that students will pursue tasks that foster increments in learning*". This is in line with our cluster motivation and its subclusters motivation and intrinsic motivation. Though, the concept of mastery or control itself was not mentioned in any of the statements. Networking strategies, a subcluster of motivation, is consistent with research by Zimmerman, Bandura, & Martinez-pons (1992), who conclude that learning strategies play an important role in academic self-motivation. More specifically, the statements in the cluster networking strategies point towards connecting to the right peers in the network. In research about creativity and innovation it is found that connecting to the right peers in a network leads to more creativity (Burt, 2004; Kratzer & Lettl, 2008).

2.5 Conclusion

This chapter presented findings of a small-scale, exploratory study, using an innovative elicitation technique called Tweetstorming; the study aimed to discover how learners perceive their personal learning in a network. Especially now that learning is increasingly using online, social technologies, a new study was needed to investigate the question at hand.

The findings will inform moderators and learners that design online, personal professional learning networks about a range of personal factors that motivate professionals to learn through networks. For example, a learner may be motivated through reciprocity (Kogut, 1989; Song, 2009) in the network (Aviv & Ravid, 2005). They want to have a *quid pro quo*; something in return for what they share in the network. For instance, in exchange for their participation and knowledge sharing, networked learners expect to receive feedback from other participants in the network. Furthermore, a personal learning network should keep a balance between an appropriate amount of information sharing and interaction in the network and a trustworthy and supportive entourage (Rusman, Van Bruggen, Cörvers, Sloep, & Koper, 2009). Future work should therefore focus on the interplay between factors that influence the interaction between networked learners.

Limitations
The results of the PLN identification session were difficult to analyse by character, as they consisted of some multiple response questions, which means that a contact could be a research collaborator and an external colleague at the same time. Also,

the response rate was very low. Further investigation with a larger group of participants is needed to allow more robust PLN identification. A further study with a larger group of participants would also allow us to aggregate the egocentric networks and compare the participants' view of their network to existing learning networks of which they are a part.

A further limitation of this study was that participants were mostly researchers already with a shared interest as evidenced by their attendance at this particular conference. Thus, the answers are likely to be in line with this type of profession. Future research should try to focus on participant groups beyond academia, in order to arrive at more general findings.

Finally, the Tweetstorm results may have been influenced by the fact that it was a brainstorm that took place via Twitter. The participants were inexperienced with such type of elicitation, which may have had its influence on the way participants expressed their statements.

Acknowledgments

We thank the respondents for their participation in the study. Furthermore, we are indebted to the experts Francis Brouns, Pia Fontana, Henry Hermans, Jo Boon, Teresa Guasch, Arnoud Evers, Bieke Schreurs, Ellen Rusman, Colin Milligan and Andrea Klaeijsen for categorizing the tweets. We thank the reviewers for their insightful comments. Finally, we thank the organisers of the PLE conference 2011 for accepting our proposal for the workshop through which we collected the data for this study.

CHAPTER 3

Factors that Influence Cooperation in Networks

When we want to know about how cooperation networks function, we could ask the network participants themselves how they perceive their learning network. However, this is only their personal perception as a practitioner. As a means to arrive at more general conclusions, in this chapter we describe an experiment with two groups of experts. They have been asked to identify their view of the set of factors that influence cooperation networks.

We built an online environment to conduct an electronic version of the Delphi method, the eDelphi. The two groups of six experts gave their view on key factors that influence cooperation networks. Group 1 was a heterogeneous group, consisting of experts in the field of network theory, behavioural game theory, social psychology and innovation and cooperation. Group 2 consisted of a more homogeneous group, comprising experts from a specific type of cooperation network: learning networks.

This chapter is based on: Sie, R.L.L., Bitter-Rijpkema, M., Stoyanov, S., Sloep, P.B. (accepted). Factors that Influence Cooperation in Networks. *Computers in Human Behavior*.

Abstract

Cooperation networks come in many forms. Innovation networks, learning networks and research networks all share the same cooperative intention, but too often they fail, as members of the network do not know which partnerships are valuable. We plan to build a support service that provides insight into the value of future cooperation, but to do so, we need to know what contributes to effective and efficient cooperation. Therefore, our main question focuses on which factors influence effective and efficient cooperation in networks. In addition to a literature review, we applied the eDelphi method to bring to light these factors. The eDelphi is a method to solicit knowledge from experts anonymously and without geographical constraints. Observations from two eDelphi rounds are reported in this chapter. The first round focused on factor generation and determined which factors influence cooperation networks and was conducted with two groups of six representative experts. Analysis of results shows that experts perceive open communication, attitude, trust, keeping to appointments and personality to be important factors that influence cooperation networks. A team of four moderators categorized the factors in a second round, resulting in four core clusters: personal characteristics, diversity, effective cooperation, and managerial aspects. A comparison with literature shows some overlap, while some factors from theory were not mentioned by the expert groups. We provide an overview of clusters identified in this study and additional factors that were missed out on.

3.1 Introduction

In everyday life, we regularly face situations in which we have to work together with others. We learn together and from others, we work together to develop new products, or we try to solve problems cooperatively. Even when we buy a product in a store, seller and buyer cooperate in favour of both. The seller earns money in order to make a living, and we get the product or service that we want. Cooperation fulfils a crucial role in our lives, for instance in the development of new products or in sharing risks (Das & Teng, 1997). When we cooperate, we connect to others, inherently constituting to a *cooperation network*.

Cooperation networks can take multiple instances. For example, innovation may take place in a cooperation network. More and more firms are now making their knowledge public in order to profit from the advancements others make with that knowledge. A recent example is Google and their Android platform. Android was released under an open source license, making it possible for others to advance Google's knowledge in the form of a mobile platform. Google in turn profits from the adoption of the platform, and starts cooperating with interesting projects, or even buys the projects. Google shares knowledge in its social network, and profits from advancements others make with that knowledge, so-called *networked innovation* or *open innovation* (Chesbrough, 2003).

Another instance of cooperation networks are learning networks. Learning networks are defined by 'non-organised groups of learners' (Berlanga et al., 2008b) that share the common intention of sharing and exchanging knowledge with the individual purpose of learning, or acquiring new skills. The nodes in the network are represented by individual learners, or even organisations that try to learn (Simon, 1991). Sharing and exchanging knowledge are the cooperative actions that define the connections between the learners. Small, temporary groups (Ad-hoc transient communities) have been proposed to guide the interpersonal relationships that are formed within learning networks by promoting sociability, trust and a sense of belonging (Berlanga et al., 2008b; Fetter, Berlanga, & Sloep, 2009).

Knowing whom to cooperate with plays a pivotal role in cooperation networks. A study among 40 managers found that one of the key determinants of effective relationships in terms of knowledge transfer and creation is valuing others and their knowledge (Cross, Parker, Prusak, & Borgatti, 2001). Selecting the right partnerships indubitably effects future cooperation (Das & Teng, 1997). Other studies show that effective cooperation within a network can boost creativity and innovation (Burt, 2004; Cassiman & Veugelers, 2006; Kratzer & Lettl, 2008; Perry-Smith, 2006). Linking to new people beyond the firm gives access to new information, assets and knowledge. New insights can be taken back to the firm to add new perspectives to current thoughts (Boland & Tenkasi, 1995).

We face a number of problems when we search for valuable peers in our network. First, as the network size increases, so does the chance of experiencing information overload (De Choudhury et al., 2008). For example, in a social network of 200 people it is considerably more difficult to distinguish valuable peers than in a network of twenty people. The people that do perceive their social network well are associated with more power, in both informal structures (friendship) and formal structures (organisation) (Krackhardt, 1990). Second, our ability to decide whom to cooperate with is bounded by cognitive limitations (Gigerenzer & Selten, 2001; Selten, 1998; Simon, 1982). If we take into account a large variety of factors that influences effective cooperation, we are not able to calculate the value of others within a reasonable time frame.

Providing insight into the value of others and their knowledge through automated software may help both individuals and teams in a number of ways. Firstly, it may give potential team members an incentive to work together. Providing team members with insight about each other may foster reciprocal action. Secondly, it helps individuals that seek for cooperation to make a satisfactory decision that would otherwise be too complex to calculate, due to cognitive limitations. Thirdly, it increases one's cognition about one's network. This has been found to correlate positively to one's power as perceived by others (Krackhardt, 1990).

To build effective and efficient software, we need to comply with two main constraints. The first constraint is the existence of a mechanism that allows us to estimate the future value of cooperation. Applying coalition theory solves the first constraint we have to comply with. Coalitions are well known in politics, where two or more parties cooperate to achieve a necessary majority in the Chamber of Deputies. Generally speaking, coalitions are temporary alliances between distinct members that cooperate. By cooperation, we mean that they share a common intention, based on individual goals (Sie, Bitter-Rijpkema, & Sloep, 2010a). Organisational teams, in essence, are cooperative in behaviour. For example, they may share the common intention of inventing a new product. They, however, do not share the same goal, that is personal growth. Game theoretic solution concepts such as the Shapley value (Hart, 1987; Shapley, 1953) and the nucleolus (Kohlberg, 1971; Schmeidler, 1969) provide an *a priori* estimation of the value of future coalitions. If we apply such calculations to teams or individuals that learn together, we may be able to determine the value of their prospective cooperation, the coalition.

The second constraint follows from the application of the above solution concepts. To provide individuals and teams with the value of potential cooperation, we need to know what factors play a part in effective cooperation. In other words, we need to know which and how factors contribute to a value for effective cooperation. Extensive literature study brought forward several factors that influence cooperation networks, such as social identity (Cheung & Lee, 2010; Keltner, Kleef, Chen, & Kraus, 2008), actor similarity (Ibarra, 1992; McPherson, Smith-Lovin, & Cook, 2001) and power (Burkhardt & Brass, 1990; Ibarra, 1993b; Swan & Scarbrough, 2005). Though, in the case of real-life intervention in human behaviour, which is inherently irrational from time to time, it is vital to have practical, in-depth expert knowledge and up-to-date knowledge about factors that influence cooperation. Hence, we employ an online, modified version of the Delphi method (Linstone & Turoff, 1975), an eDelphi (Bitter-Rijpkema, Martens, & Jochems, 2002), to elicit that knowledge.

The Delphi method aims to solicit information and ideas from a panel of experts about a specific subject through a series of opinion expression. The Delphi has been recognised as one of the most effective approaches for getting a consensual agreement among experts on particular issues (Davis & Alexander, 2009; Hasson, Keeney, & McKenna, 2000; Kennedy, 2004; Linstone & Turoff, 1975; McKenna, 1994). Because domain experts are likely to be well informed about the latest technologies and their adoption, the Delphi method is often used to identify trends (Davis & Alexander, 2009; Milkovich, Annoni, & Mahoney, 1972; O'Neill, Osborn, Hulme, Lorenzoni, & Watkinson, 2008; Rice, 2009). The Delphi has a number of advantages. First, there is no need for experts to discuss face-to-face, as the questionnaires are sent to participants. Originally, the Delphi was sent by mail, but recent approaches make use of online versions of the Delphi (Distler et al., 2008). Second, as there is no need to discuss face-to-face, the Delphi may be conducted anonymously. Alternatives such as brainstorming (Osborn, 1954) or focus groups (Merton, 1984) cannot be conducted anonymously, as participants meet face-to-face. Third, the discussion between participants can change the opinions, and they have the opportunity to change them throughout the process as multiple questionnaires or 'rounds' are conducted. Brainstorming, for instance, is focused on generating as many ideas as possible, and thus does not allow participants to criticize other's ideas during the process.

The original Delphi was sent by paper mail and comprised a series of questionnaires, in which opinions were fed back to participants in a next questionnaire. In this way, agreement among participants could be reached. Today's technology (forums, chat, wikis) allows online discussion; therefore we conduct eDelphi, an electronic version of the Delphi, in a tailored online environment. Also, our aim is slightly different. We do not search for consensual agreement, rather we search for complementary knowledge that experts may have about cooperation networks. The eDelphi comprises two rounds in which factors are generated, rated and clustered. This chapter reports on the results and findings of two rounds of the eDelphi: factor generation stage performed by participants, and the factor clustering stage performed by a team of moderators. The focus is on the following question, which will be presented at the very start of this eDelphi session: *What factors influence cooperation networks?*

The structure of this chapter is as follows. In Section 3.2, we lay out our research methodology, which includes a description of the eDelphi method and the procedure. Section 3.3 presents the results of each round separately, as round one was conducted with two panels of experts, and round two was conducted with a team of moderators. We will discuss the results in Section 3.4 and draw our conclusions in Section 3.5.

3.2 Method

3.2.1 The eDelphi method

To identify the factors that influence cooperation networks, we applied the eDelphi, a modified version of the Delphi method. It took place on the Internet during a four-week period in April and May 2011, via an advanced, tested environment. An introductory statement welcomed the participants to the environment. The introductory statement provided the participants with the main question *What factors influence cooperation networks?*, and a context description to clarify the main question. Special attention was given to the context description. We were aware of the fact that too much information could bias the participants. Therefore, we decided to have a short, but satisfying description of a cooperation network, and a real life example, without specifically mentioning factors that influence, or characteristics of a cooperation network. Next to the context description, we provided Twitter, Delicious and Google News feeds that contained the words 'cooperation' and 'network' (Figure 3.1b) to provide a better understanding of the concepts cooperation and network. It also provided the necessary additional information to sufficiently create a context for the question at hand, without constraining the participants to think in a certain direction.

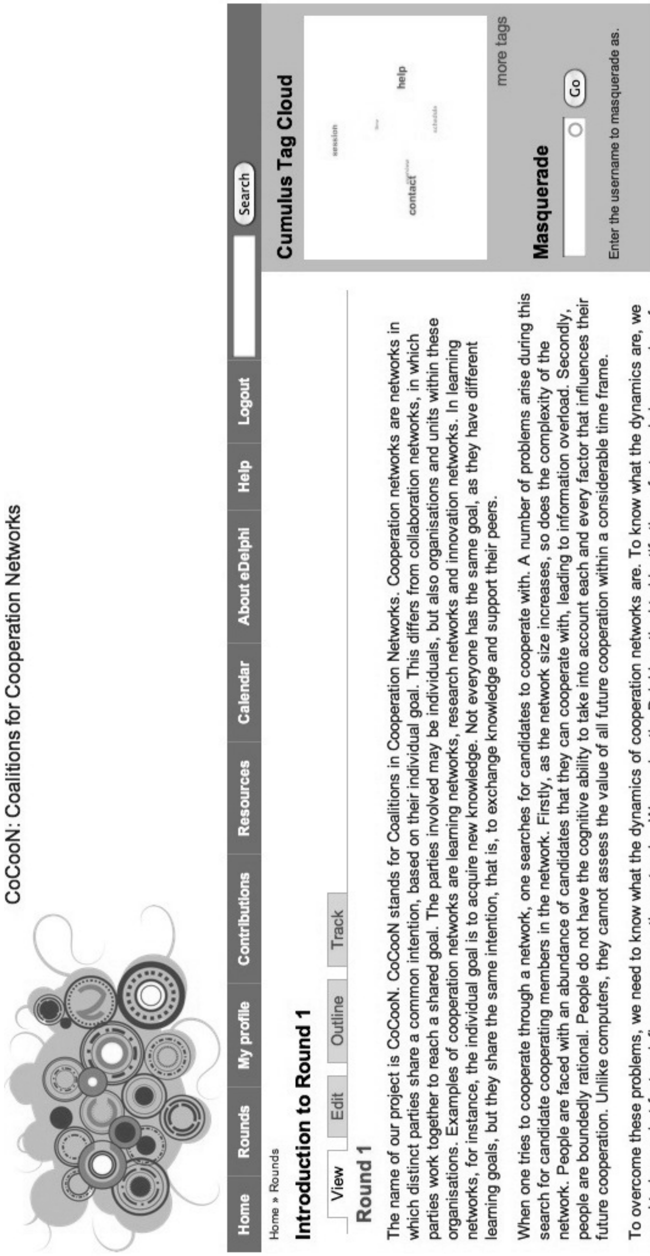


Figure 3.1a. Top half of a screenshot of the eDelphi environment. The main content describes the context.

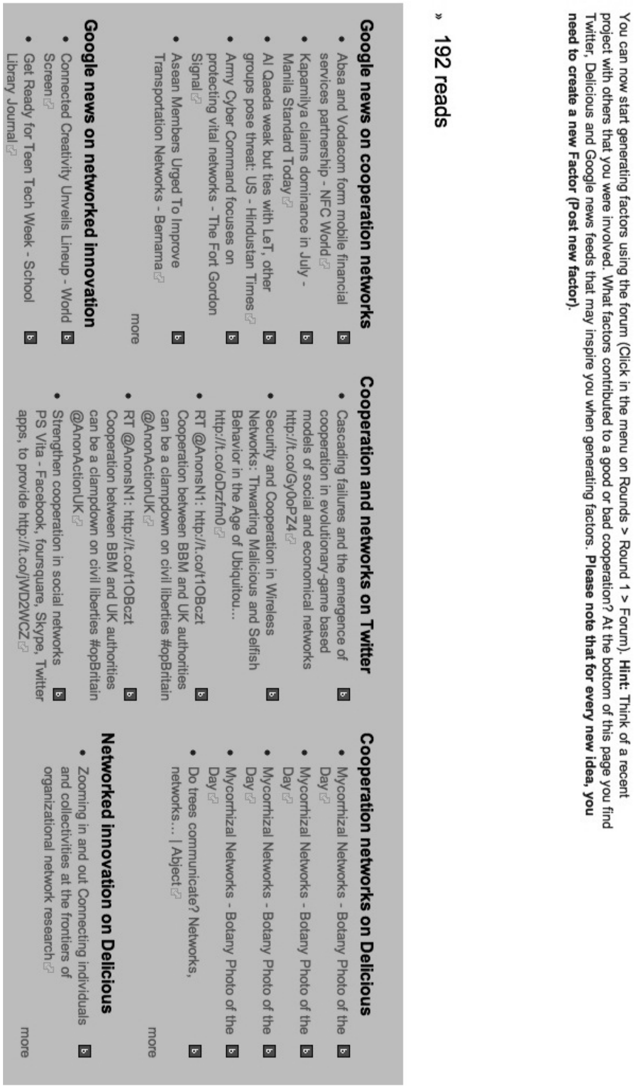


Figure 3.1b. Bottom half of a screenshot of the eDelphi environment. Additional feeds from Twitter, Delicious and Google News provide the necessary context.

During the first round that took four weeks in April and May 2011, experts could articulate factors via forum posts. Factors could be discussed by leaving a reply on the individual page of a posted factor. The factors were quasi-anonymous, as the facilitator could see who contributed the factors. This was especially important in case one or more participants would become inactive during the process. Participants could be addressed personally to state that they have been inactive for a while. Also, in case of inactiveness, it was easier to discover why participants failed to be active in the environment.

The factor generation round of the eDelphi is a round of opinion expression, perspective taking and idea generation. Therefore, it is important to generate factors from a wide range of perspectives. We must be cautious, though, not to overlook certain specific factors. We therefore choose to have two groups of experts to cover both general and specific factors: one group of experts that represented expert from a broad area of expertises that are relevant to cooperation networks, and a second group of experts from a specific instance of cooperation networks, namely, learning networks. Naturally, we decided not to merge the two groups, as this may have resulted in the generation of general factors.

After generation of factors, the participants were asked to state how important they found the factors. On the individual page of a factor, ratings on a scale of one to five stars could be assigned; one star meant ‘not important’, five stars meant ‘very important’. We explicitly did not ask participants to rate each and every factor, as this could increase workload drastically as the number of factors increased. The ratings were conducive to a correct interpretation by the moderator team that made a summary of the Delphi session. Voting allows participants to make a decision which opinions to accept or reject. It is relatively quick, but restricted, as it does not care for gradual expression of participants’ preferences for opinions. Ratings allowed the participants to express for every opinion to what extent this was preferred. Regularly, the facilitator would feed back the factors that were generated, to trigger new discussion and factors.

During the second round that took a week, a team of moderators analysed the factors that were generated. In the development of a system model that simulates and recommends optimal future cooperation it is important to have a set of core clusters, rather than a large set of factors that act as variables. It is commonly acknowledged that a system that uses more variables to represent reality is also more prone to errors. The factors were fed into the WebSort.net (http://websort.net) clustering environment. WebSort provides a variety of data aggregation (e.g. items vs. items, items vs. categories) and visualisation opportunities (e.g. tree structure, tables). Moderators could add factors to self-defined clusters with self-defined names. On purpose, we chose not to use predefined cluster names, to prevent bias from the researchers. Subsequently, overlap between the clustering of the moderators was computed using agglomerative hierarchical cluster analysis.

To goad correct interpretation of the eDelphi, we include an overview of the workflow in Figure 3.2.

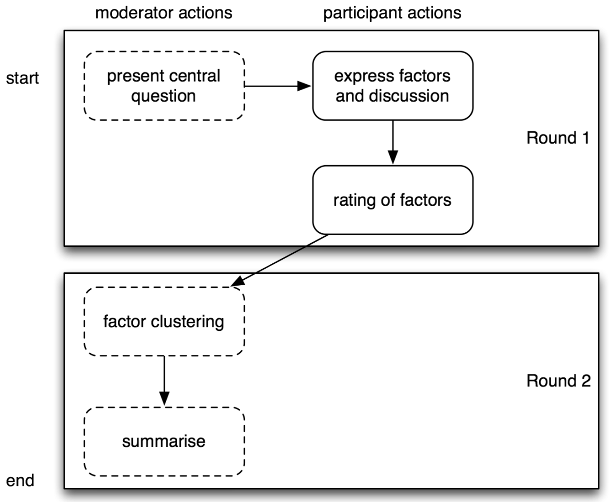


Figure 3.2. Overview of actions for the eDelphi. Boxes surrounded with a dotted line (left) are moderator actions. Normal boxes (right) are participant actions.

3.2.2 Participants

Group 1 consisted of recognized senior professionals with knowledge and experience in the following knowledge areas: 1) Network theory, 2) (Behavioural) game theory, 3) social psychology, and 4) innovation/ cooperation. By senior professionals, we mean academic staff that has a doctorate or higher, or business professionals with five or more years working experience in one of the aforementioned areas. Table 3.1 shows the knowledge areas the experts are working in. In total, group 1 consisted of six experts.

Table 3.1. Main expertise of experts in group 1.

Expert	**Network theory**	**(Behavioral) Game theory**	**Social psychology**	**Innovation/ cooperation**
1	x			
2	x	x	x	x
3	x			x
4	x			
5	x			x
6	x		x	x

Group 2 consisted of six experts in the field of learning networks. The learning networks experts have more in-depth and practical knowledge. Besides, they are more likely to agree on the more specific factors, as they have the same experience with learning networks.

Larger sample sizes (up to twelve participants) have been reported to generate more and better ideas (Gallupe et al., 1992). Though, after a certain threshold, groups become saturated; there seems to be no difference between, for instance, eight and forty-eight participants with respect to the number of relevant ideas generated (Aiken, Krosp, Shiran, & Martin, 1994). Several studies on cost-effectiveness in usability studies support this by claiming small sample sizes (Turner, Lewis, & Nielsen, 2006; Virzi, 1992). Having said that, we think that a total sample size of twelve is sufficient for generating factors to be clustered in core groups of factors by expert moderators.

3.2.3 Data Collection and Analysis

The factor generation primarily resulted in two sets of factors, each by one of the expert groups. Analysis of the resulting factors informed us about the activity of the participants. We could also distinguish between the groups based on the character of their output. Unlike Hasson, Keeney and McKenna (Hasson, Keeney, & McKenna, 2000), there was no need to discover factors and discussion, as they showed up in the forum when they were posted. The factors and discussions could be posted by participants directly, without any interference of the facilitator or moderator team. To do so, the participants received a personal login to access the eDelphi environment.

Next to factors generation, we asked the participant groups to rate how important they found the factors, based on a five-star scale. To rate a factor, participants would click on a factor to visit its page, and a five-star rating could be given by clicking on the appropriate number of stars. In case too many factors were generated in round 1, this could be used to make a selection of factors.

As said earlier, the results of the factor clustering were analysed using agglomerative hierarchical clustering. As we use an item-item similarity matrix to analyse the similarity between factors/clusters, we use agglomerative hierarchical cluster analysis, which starts with all factors in separate clusters. In several phases, clusters are merged based on their similarity, until the appropriate number of clusters is reached. If the average distance between two clusters was small, the clusters were merged. The similarity of factors was based on the number of times two factors co-occurred in categories defined by the four members of the moderator team. For instance, if moderator 1 put factor A and B in one category and moderator 2 put factor A and B in one category, then the similarity between factor A and B increased. This similarity measure was category-independent, which was helpful since moderators could name their own categories.

3.3 Results

3.3.1 Round 1: factor generation and rating

In round 1, the participants generated a total of 33 factors. Group 1 generated 13 distinct factors, and group 2 generated 21 distinct factors. As expected, the factors were different, and only one factor, trust, overlapped. After the factor generation, participants were asked to give a rating on a five-star scale to state how important they perceived factors to be. Table 3.2 presents an overview of the factors, their average rating (second column) and the number of ratings they received (third column).

Table 3.2. Factors generated in round 1, sorted per group and perceived importance.

Group 1	**Average rating (stars)**	**No. of ratings**
Social capital	5	1
Trust	4.5	4
Leadership	4	1
Shared goals	3.5	2
Managing cultural differences	3	2
Consciousness	3	1
Knowledgeable intermediary	3	1
Fun, good working spirit	2.5	2
Complementary knowledge	2	2
Recognizing and creating win-win situations	0	0
Clear contracts	0	0
Managing diversity	0	0
Interdependency	0	0
Group 2		
Open communication	4.75	4
Attitude	4.2	5
Trust	4.2	5
Keeping to appointments	4	4
Personality	4	5
Openness in planning	3.8	5
Work ethics	3.75	4
Humor	3.75	4
Transparency	3.75	4
Mutual respect	3.5	4
Honesty	3.4	5
Drive	3.25	4
Joint interests	3	3
Personal goals	2.8	5
Passion	2.75	4
Convenience	2.67	3
Boundaries	2.5	2
Security	2.5	4
Responsiveness	2.4	5
Diversity	2.4	5
Quality assurance	2	4

The factors in Table 3.2 are sorted according to their perceived importance. The participants in group 1 rated social capital, trust, leadership, shared goals and managing cultural differences to be most important. However, the average number of ratings per participant (5.33, Table 3.2) and response rate (0.5, Table 3.3) show that the activity for group 1 is lower, which suggests that they may be less reliable than the ratings of group 2. Group 2 perceived open communication, attitude, trust, keeping to appointments and personality to be most important. The average number of ratings per participants (17.6) and the response rate (0.83) suggest that the ratings for group 2 are more reliable.

Table 3.3. Summary of the factors and ratings generated.

	Group 1	**Group 2**
Participants		
N	6	6
Response rate (factor generation)	0.83	0.83
Response rate (factor rating)	0.5	0.83
Average no. factors per participant	2.6*	4.2*
Average no. ratings per participant	5.33*	17.6*
Factors		
N	13	21
Minimum factors	0	0
Maximum factors	6	6
Mean	2.33	3.5
Std. deviation	2.16	2.43
Ratings		
N	16	88
Min. value	2	1
Max. value	5	5
Mean rating value	3.44	3.32
Std. deviation	1.09	1.11

* based on the response rate

Furthermore, Table 3.3 includes some statistics about the factors and ratings, respectively. We see that the minimum number of factors that were generated by either groups is zero. This means that there was at least one person per group who was inactive during the generation of factors.

3.3.2 Round 2: factor clustering

The moderator team aggregated and grouped the factors together. From a methodological perspective, the clustering and rating are two different types of analysis of the data. Rating determines the popularity of the factors as perceived by the participants. Clustering combines factors share meaning into groups. Thus, the end product of factor rating is a ranked list of popular factors, whereas clustering results in multiple groups of factors that share a meaning. Analysis (Independent Samples Mann-Whitney U test) shows that the two groups do not differ significantly (.25). From a practical perspective, there were two other reasons to use all factors of both groups for the clustering: 1) there was no need to pre-select factors from either group, as few factors were identified, and 2) group 1 generated considerably less factors and less ratings than group 2, which makes it difficult, if not impossible, to determine which factors should be taken to the factor clustering round.

The items were grouped in categories by a team of four expert moderators in the fields of social networks, learning, interpersonal relationships, innovation and creativity. Table 3.4 shows the categorizations for each of the factors. The values represent the percentage of the moderators that placed the factor in that category. For instance, ‘humor’ was placed in the category ‘Emotion and Mode’ 25% of the cases, which translated to one moderator. As Table 3.4 shows, nearly no overlap in categorization was found. Only ‘social capital’ was placed in the category social capital’ in 50 percent of the cases. The reason for this is clear and expected; the moderators could define the names for the categories themselves.

Table 3.4. The percentage of times a factor was placed in a self-defined category. Rows represent factors, and columns represent the categories.

This table shows the % of times each item was placed in each group	**attributes for effective Cocoon collaboration**	**Characteristics of interpersonal relationships**	**Cocoon network characteristics**	**collaboration strategies**	**communication**	**context**	**Contextual characteristics**	**Emotion and Mode**	**group composition**	**Knowledge**	**management**	**managing diversity**	**motivation**	**mutual goals**	**openess and transperency**	**Operational, strategical and managerial working skills**	**personal attributes for COCOON collaboration**	**Personal characteristics**	**Personality**	**quality**	**reciprocity**	**security**	**social capital**	**successful cooperation**	**Trust and respect**	**values**
consciousness																	25	25	25					25		
attitude													25				25	25	25							
personality									25								25	25	25							
personal goals													25			25	25		25							
passion								25					25				25	25								
drive								25					25				25	25								
humor								25									25	25								25
fun, good working spirit								25					25			25	25									
honesty																	25	25							25	25
work ethics																25	25								25	25
diversity			25				25		25			25														
clear contracts	25				25		25								25											
open communication	25	25			25										25											
transparency	25					25	25								25											
openness in planning	25	25									25				25											
responsiveness		25			25										25		25									
social capital	25	25																					50			
complementary knowledge	25	25							25	25																
convenience		25											25				25						25			
shared goals	25						25							25									25			
security	25						25															25	25			
quality assurance	25						25													25			25			
interdependency		25	25																				25	25		
joint interests		25	25											25									25			
knowledgeable intermediary	25						25			25	25															
recognizing and creating win-win situations				25									25			25							25			
leadership	25										25	25				25										
managing cultural differences	25										25	25				25										
managing diversity	25										25	25				25										
keeping to appointments	25														25	25					25					
boundaries				25		25	25																25			
trust		25															25							25	25	
mutual respect		25															25				25				25	

Although nearly no overlap could be found, it is still possible to see that factors were placed in the same category even if the category's name was not the same. We used this to identify the similarity between factors and their categorizations. Based on agglomerative hierarchical cluster analysis, four core clusters could be identified, as depicted in Figure 3.3. The aggregated factors in one cluster are shown by a grey rectangle. As diversity is the only factor in that cluster, it shows no rectangle.

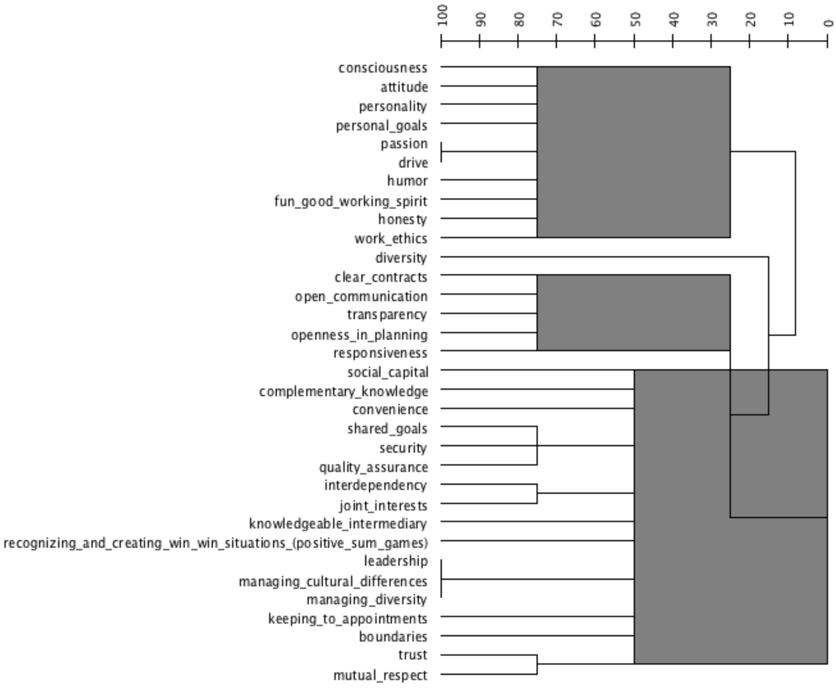


Figure 3.3. Clusters identified using agglomerative hierarchical cluster analysis.

The first cluster is mainly about personality and motivation and consists of ten factors, namely: consciousness, attitude, personality, personal goals, passion, drive, humor, fun and good working spirit, honesty and work ethics. When we look at the factors per group in Table 3.2, we see that eight of these factors were named by group 2, and two were named by group 1. Even though group 2 generated more factors in general, group 2 seems to put more emphasis on this cluster than group 1.

The second cluster that came forward using the described method was diversity, containing only one factor. This is caused by the fact that, apparently, there was no convergence in the way the moderators clustered this factor, which resulted in it being a cluster itself. There may be a number of reasons for this. First, diversity may be a cluster on its own, which is very unlikely. Second, diversity is a cluster, but no other factors that belong in that cluster were named; too little factors were named. Third, the moderator team showed too little overlap; this may be due to the self-defined category names.

The third cluster is about effective cooperation, and contains factors such as clear contracts, open communication, transparency, openness in planning and responsiveness. Again, only one out of five factors was named by group 1. This suggests that group 2 focused more on the effectiveness of cooperation.

The fourth cluster is about management and interpersonal relationships. It includes social capital, complementary knowledge, convenience, shared goals, security, quality assurance, interdependency, joint interests, knowledgeable intermediary, recognizing and creating win-win situations, leadership, managing cultural differences, managing diversity, keeping to appointments, boundaries, trust and mutual respect. Here, both groups have generated nine factors out of seventeen (trust overlaps). Given the number of factors generated, group 1 seems to have put more emphasis here.

3.4 Discussion

The main objective of this chapter was to find additional factors that were not mentioned in theory, due to their practical nature. We report the process and results of the eDelphi method that we used. It is an important step towards the development of a service that recommends valuable peers for cooperation in a network. The computation of valuable peers is based on factors that influence cooperation in a network. Therefore, we investigated the following main question: *Which factors influence cooperation networks?*

The factor clustering round produced four core clusters. When we take a close look at the categories the factors are placed in (Table 3.4), we see that the factors in cluster one are about personal characteristics. This is in accordance with personality as pointed out by Brass, Galaskiewicz, Greve and Tsai (2004). The second cluster, diversity, is underlined by various studies as a key factor for knowledge sharing (Berendt & Kralisch, 2007) and perspective taking (Boland & Tenkasi, 1995). The third cluster describes effective cooperation. It is important to effectively cooperate, as it is a core activity in cooperation networks such as interfirm alliances (Das & Teng, 1997). The fourth cluster is about the managerial aspects of cooperation networks. Schreiner, Kale and Corsten (2009) note that the capability to manage cooperation is key to its success. They mention motivation (identifying potential benefits), choosing the right partners, effective communication, and developing strong ties as key management activities. In our view, these are in agreement with the factors joint interest, shared goals, security, trust, mutual respect and interdependency that are identified in this study.

If we compare the factors and clusters to literature, we see that a number of factors were not mentioned. Perhaps this is due to the nature of the discussion or the context that was given, but little factors were named that influence cooperation networks badly. For instance, accountability (Jensen & Roy, 2008; Tetlock,P. E., 1992) and social loafing (Chidambaram & Tung, 2005; Latane, Williams, & Harkins, 1979; Liden, Wayne, Jaworski, & Bennett, 2004) were not mentioned. Also, factors concerned with the value future cooperation partners, such as power (Keltner et al., 2008), status and reputation (Jensen & Roy, 2008), and actor similarity (Ibarra, 1992; McPherson et al., 2001) were not mentioned. Decision-making flaws such as escalation of commitment (Shubik, 1971), risk or loss aversion (McCarter, Rockmann, & Northcraft, 2009) and groupthink (Janis, 1982) also remained unidentified. Table 3.5 shows an overview of clusters found in this study and their basis in literature, and additional factors from literature that were missed out on.

Table 3.5. Factors identified in this study, and factors that were mentioned in literature.

Factors/clusters	**Literature**
Current study	
Personal characteristics	(Brass et al., 2004)
Diversity	(Berendt & Kralisch, 2007; Boland & Tenkasi, 1995)
Effective cooperation	(Das & Teng, 1997)
Managerial aspects	(Schreiner et al., 2009)
Additional from literature	
Accountability	(Jensen & Roy, 2008; Tetlock,P. E., 1992)
Social loafing	(Chidambaram & Tung, 2005; Latané et al., 1979; Liden et al., 2004)
Power	(Keltner,Dacher, 2008)
Status and reputation	(Jensen & Roy, 2008)
Actor similarity	(Ibarra, 1992)
Escalation of commitment	(Shubik, 1971)
Risk/loss aversion	(McCarter et al., 2009)
Groupthink	(Janis, 1982)

Group 2 has generated considerably more factors and ratings, which makes their ratings more reliable. The factors that are perceived most important are open communication, attitude, trust, keeping to appointments and personality. Jarvenpaa and Leidner (1998) show that predictable, thus good communication is key to trust within global virtual teams. Furthermore, they state that teams that end a project with high levels of trust focus on procedures and tasks and show professional relationships. This may be in line with keeping to appointments, although on a more abstract level. Brass *et al.* (2004) acknowledge the existence of attitude in interpersonal networks, but rather see this as a consequence of cooperation in a network. Brass *et al.* highlight a number of factors that foster interpersonal networks: actor similarity, personality, proximity and organisational structure, and environmental factors. Personality is in line with the findings of our study. Though, the factors found here are subject to the context of the participants. The participants of group 2 work in a specific instance of cooperation networks, learning networks, and these factors may be only relevant for learning networks.

The interpretation of the results poses some methodological considerations. The eDelphi was conducted solely online and the design of the environment made it possible for participants to contribute anonymously. Being anonymous has a number of advantages such as no emergence of a hierarchy, which may be very important when you want to discover the real opinions of people. Anonymity also has some drawbacks, as people cannot be accounted for their lack of contributions. We therefore chose to let the participants be quasi-anonymous; They were anonymous among the group, but not to the facilitator. The facilitator could remind them to contribute.

Despite numerous attempts to regenerate the discussion and generation of factors, the experts in group 1 remained very inactive. Distler *et al.* (2008) state that a lower response rate may also be due to the fact that participants were not member of a pre-existing expert group. Some studies provide current information on the subject in the first round to be rated. The advantage of such a round is that participants have a clear picture of the context of the subject right from the start. A disadvantage may be that participants will be subject to bias. We think that the low response rate of group 1 during factor generation may be due to the absence of an extensive description of the context of the problem.

A challenge lies in the optimisation of the eDelphi process. When using a diverse group (group 1), the activity for round 1 was very low. Factors generated seemed to be more general and focused on the managerial aspect of cooperation networks. Possible improvements may be publishing a pre-study survey on the subject, to provide a clearer context for factor generation and discussion. Also, accountability and the number of facilitator interrupts may be increased to raise activity among the participants.

3.6 Conclusion

In this chapter, we presented an online expert Delphi that inquired experts about factors that influence cooperation networks. We reported two rounds of the eDelphi: 1) factor generation and rating, and 2) factor clustering. Key factors as perceived by experts include effective communication and trust formation, attitude, process and task focus and personality. Factor clustering by a team of moderators and agglomerative hierarchical cluster analysis resulted in four core clusters of factors. These clusters describe personal characteristics, diversity, effective cooperation and management and interpersonal relationship. The diverse group of experts (group 1) focused on the managerial aspects of cooperation networks. The experts specialised in learning networks (group 2), a specific instance of cooperation networks, rather focused on effective cooperation and personal characteristics.

Furthermore, a comparison with literature showed that there is overlap in both theoretical and practical knowledge, but that some factors remained unidentified by the expert groups, such as status, power, reputation, accountability and social loafing. This may be due to the character of the discussion or the context description that was given in advance. This may need some extra investigation, but on the other hand, we contend that the sum of theoretical and practical knowledge has given us a well-elaborated picture of factors that influence cooperation networks.

Now that we have laid a proper theoretical and practical foundation of factors that influence cooperation networks, we proceed with further steps in the design and implementation of the system we plan to develop. Roughly speaking, the following steps in the design of our system are: 1) definition of a system model or architecture (design), 2) a simulation of cooperation networks (validation), and 3) recommendation of future valuable peers for cooperation (implementation).

CHAPTER 4

What's in it for me? Recommendation of Peers in Networked Innovation

One of the aims of this thesis is to support individuals in finding the right peers for cooperation. From a methodological perspective, this requires an intervention study to test proposed support tool with subjects. However, intervention studies may consume a lot of time in terms of preparation and performing the intervention itself. A simulation in advance gives insight into how certain factors influence one another, and how they influence the subjects. Being informed by a literature review and the two studies of factors that influence cooperation networks (Chapters 2 and 3), we implemented a simulation of how people form connections in innovation networks, which are an instance of cooperation networks.

This chapter is published as: Sie, R. L. L., Bitter-Rijpkema, M., & Sloep, P. B. (2011). What's in it for me? Recommendation of Peers in Networked Innovation. *Journal of Universal Computer Science*, *17*(12), 1659-1672.

Abstract
Several studies have shown that connecting to people in other networks foster creativity and innovation. However, it is often difficult to tell what the prospective value of such alliances is. Cooperative game theory offers an a priori estimation of the value of future collaborations. We present an agent-based social simulation approach to recommending valuable peers in networked innovation. Results indicate that power as such does not lead to a winning coalition in networked innovation. The recommendation proved to be successful for low-strength agents, which connected to high-strength agents in their network. Future work includes tests in real-life and other recommendation strategies.

4.1 Introduction

Several studies argue that groups are more innovative than individuals (Paulus & Yang, 2000; Paulus, 2003). Individuals by themselves do not possess all the knowledge that is needed for innovation, for innovation to be successful it requires networked interactions (Downes, 2003). That is, knowledge has become diffused, as Henry Chesbrough (2006) emphasises. He argues that, to keep up with today's dynamically changing environment, firms need to adopt *open innovation*. It occurs as a result of opening up, or freely distributing knowledge. Thereby, a firm profits from 1) the advancements others make with that knowledge and 2) complementary knowledge that lies beyond the borders of the firm. This is consistent with earlier work by Barnard (1968) and Simon (1991) that firms cannot rely on their own internal knowledge to flourish. Viewed from a collaborative learning perspective, Yazici (2005) found that a collaborative learning style influences team performance positively. Cassiman and Veugelers (2006) proved that complementary knowledge present in an R&D's social network may significantly boost new product development. This network perspective on creativity and innovation is highlighted by a number of studies: Kratzer and Lettl (2008) concluded that people that are on the edge of two social networks, so-called 'lead users', tend to be more creative than others in their network, as they are more informed. Ronald Burt (2004) uses the term 'brokerage' to denote the same phenomenon. Perry-Smith (2006) stresses the importance of a central network position and weak ties beyond the borders of the firm in order to be more creative.

Even though the network perspective to creativity and innovation is a promising way of dealing with knowledge, it is not without problems. While people engage in knowledge sharing activities in their network, they need to be aware of which people are most valuable to them. Psychological research points out various decision-making problems, such as bounded rationality (Simon, 1982): Due to cognitive limitations and incomplete knowledge, people are not capable of computing probability in a reliable way, being 'boundedly rational'. In networked innovation, bounded rationality is encountered in a similar way. While searching for valuable peers, one is faced with an abundance of peers to connect to (information overload / incomplete knowledge) and our minds lack a proper metric for assessing the value of peers (cognitive limitations).

The human mind is complex and it is thus challenging to model its cognitive abilities. Cooperative game theory addresses this complexity by assuming human beings – players – to behave rationally. Cooperative game theory describes decision making about cooperation in a game. It enables one to make an a priori estimate of the value of cooperation. Such an estimate strengthens one's cognition of the network, which is found to positively correlate to power as perceived by others (Krackhardt, 1990). Agent simulations are an often used approach to model players in a network, using game theoretic considerations. Previous studies that simulated creativity and innovation include the use of computer simulation (Phelan, 2002), system dynamics (Wu, Kefan, Hua, Shi, & Olson, 2002), agent-based simulation (Schwarz & Ernst, 2009; Albino, Carbonara, & Giannoccaro, 2006; Ma & Nakamori, 2005) and swarm-based simulation (Bhattacharyya & Ohlsson, 2010).

In this chapter, we model observations from literature to simulate behaviour in networked innovation. Recommendations are generated to inform agents about the value of peer agents. In Section 4.2, we provide the underlying theory necessary for understanding the proposed simulation method, which is described in Section 4.3. Section 4.4 comprises the results of our simulation, which we will discuss in Section 4.5. Future work is discussed in Section 4.6.

4.2 Theoretical Background

4.2.1 Game Theory

A 'game' in the sense of game theory is a situation in which one or more players use strategies to optimise their reward. Rules of play identify the character of the game and players have to comply with these rules. Games such as Chess are played for fun, but more serious and realistic games are played as well. In daily life, games (in the game-theoretic sense) are played every day and everywhere. Though, many of us are not aware that they are playing a game. On eBay, buyers that bid for a product play a game against each other and the seller of that product. In labour negotiation, a game is played between future employee and future employer. Each game has one or more players. Players comply with a set of rules that define the game. Players strive to win (or optimise their outcome), and this may result in competing (non-cooperative) play against others, or cooperative play with others. To optimise the outcome of a game, a player follows certain strategies, or heuristics to win a game. Such strategies often include an estimate of a game's prospective reward, which is called the *expected utility*. A player can win everything, like a product in the auctioning game in the eBay example, but this means the other players lose. A player can negotiate an outcome, like in contract negotiation. When a game of Chess is played, a player may win (+1), draw (+0) or lose (-1). Chess is a *zero-sum* game. A game is said to be zero-sum if the sum of wins (+1) and losses (-1) of all players equals zero. Akin to zero-sum games, a *constant-sum* game is a game in which the sum of all wins and losses equals a constant. The bidding game on eBay is a constant-sum game, as one player wins and pays for a product and the other players lose and pay nothing. The constant sum in this game equals the price of the product. The reward that you receive after playing a game is called the *payoff*. Players try to rationalise what other players are about to do, to maximise their payoff.

4.2.1.1 Coalitions

For clarifying purposes, we have to distinguish between cooperation, collaboration and coordination. When people decide to work together, based on their individual goals, we speak of *cooperation* (Axelrod & Hamilton, 1981). When people work together, based on common goals, we speak of *collaboration*. When people agree to perform the same actions (interactional synchrony), we speak of *coordination* (Arrow, McGrath, & Berdahl, 2000). When people cooperate temporarily and coordinate their actions, a coalition is formed. In other words, a coalition is a temporary alliance in which players share a common intention. It is, however, based on individual interest, or goals (Cyert & March, 2005). A labour contract can be seen as a coalition. Employee and employer agree to a common intention, that is, work for the company, but they have individual goals: the employer wants to make profit, and the employee wants to earn a living. Coalitions are often formed in games in which the payoff can be divided among members of a coalition. If a payoff can be divided, or transferred without costs, we may speak of *transferrable utility*. What characterises a cooperative game with transferrable utility, is that it is often more profitable to form a coalition and share the payoff, than to go it alone and most likely receive less or nothing.

Shapley Value

The Shapley value (Shapley, 1953; Hart, 1987) was designed by Lloyd Shapley in 1953 to evenly distribute the payoff in a game with transferrable utility among members of a coalition. The Shapley value is calculated by measuring the strength of a coalition, minus the strength of its subcoalitions. Subcoalitions may consist of multiple persons, but one-person and zero-person coalitions may also be identified.

4.2.2 Agent-based Social Simulation

Agent-based social simulation is a way to understand certain social phenomena through simulations of agent societies. According to Davidsson (2002), this field can be best characterised by the intersection of social science, computer simulation, and agent-based computing. Social science is the study of social phenomena done in a variety of research areas, such as social psychology, biology and economics. Computer simulation is a field in computer science that is used to study social events. The aim is to predict future behaviour of such a social event. Agent-based computing is also a field in computer science and it includes intelligent agents and multi-agent systems. Agents are computer programs, that are supposed to act autonomously, pro-actively, reactively, and socially able (Wooldridge, 1998). In multi-agent systems, agents interact with each other, often to solve a (divisible) problem or to observe the agents' behaviour.

4.3 Simulation method

4.3.1 Simulation Model

Below, we provide the model used for simulation of coalitions in networked innovation (Figure 4.1). This model may be regarded as the internal reasoning structure of an agent.

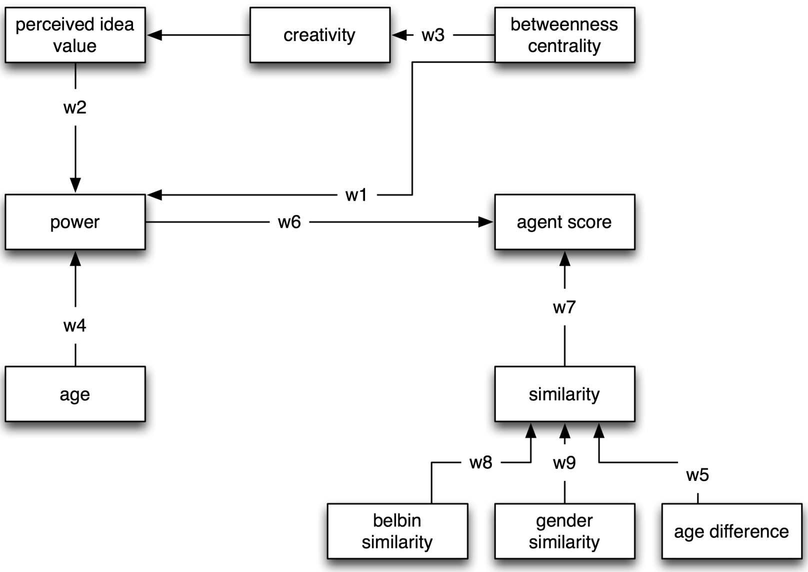


Figure 4.1. The simulation model; for a detailed description, see text.

Two factors are highly influential for the formation of coalitions: 1) *power* and 2) *similarity* between people (homophily). These two directly contribute to an agent's score for each of the agents in our model. An agent's score determines the likelihood that an agent is interested in forming a coalition with another agent. There are seven factors that indirectly, through the two central factors, contribute to an agent's score.

From Social Network Analysis Theory (Wasserman & Faust, 1994), we choose to use the concept of betweenness centrality to express someone's position in the organisation. Betweenness centrality is a measure of how dependent others are one a target node in a network. It is computed by the number of shortest paths that pass through a node, as a proportion of all shortest paths possible. In our case, betweenness centrality measures how dependent people are on one another if they want to connect. People cannot form a coalition if there is no path that connects them. If an agent possesses high betweenness centrality, agents very likely have to pass him to reach any one person in the network. Betweenness centrality influences a number of factors. Firstly, Kratzer and Lettl (2008) found that 'lead users', people that are on the edge of two networks, are more likely to be creative than others. Tsai and Ghoshal (1998) underscore this by reporting that social interaction (often viewed as degree centrality) and resource exchange were positively correlated to product innovations. Kraatz (1998) extends this view by emphasising that interorganisational ties may advance social learning, thereby contributing to organisational growth. Secondly, various studies report that people that are more central are found to be more powerful (Perry-Smith, 2006; Krackhardt, 1990; Ibarra, 1992; 1993a; Brass, 1984).

Power is also influenced by age and the perceived value of an idea. Age is reported to correlate positively with power (Burkhardt & Brass, 1990). Klein and Sorra (1996) suggest that 'innovation-values fit', the extent to which an innovation (idea) fits the perceiver's values, influences . In our model this is represented by the perceived value of an idea.

Herminia Ibarra (1992) reports that similar people (homophily) are more likely to form support and friendship relationships. This is emphasised by McPherson et al. (2001). They distinguish between various types of homophily, such as age and gender. For our model, we use age, gender and personality to express similarity.

4.3 Agent Characteristics

Age is represented as a random value between 15 and 65, the so-called 'working age' of people. Gender is represented as a random value of 0 (female) or 1 (male). Personality is difficult to represent. Multi-attribute personality scores such as the Big Five personality traits have been considered, but for the time being, we choose to use the Belbin Team Roles (Belbin & Belbin, 1996). The nine Belbin profiles express the role of a person within a team. Use of these predefined team roles eases the computation of similarity.

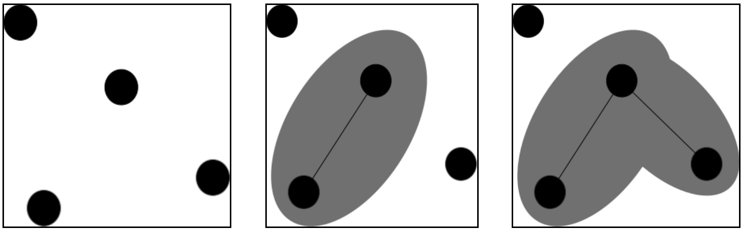
Agents have a power attribute, which corresponds to their power in the model. Agents' ultimate score is influenced by both their power and their similarity to other agents.

4.3.1 Network Characteristics

Akin to common networks, the network of innovators we model consists of nodes and links. Every node represents a person. Bilateral links between these nodes denote professional relationships between these persons. Combinations of links make paths through which people can be reached. A network is defined by its size (the number of agents/ people), its density (the number of links between people as a proportion of all possible links) and the path length. We use shortest paths between people to compute betweenness centrality.

4.3.2 Coalitions

If two agents decide to cooperate, they form a dyadic connection. Afterwards, all dyadic connections that overlap are gathered, thereby forming paths between multiple agents. These paths of accumulated dyad connections form a subnetwork within the whole network of agents. Such a subnetwork of cooperating agents we have called a coalition (see Figure 4.2).



2a 2b 2c

Figure 4.2. Evolution of a coalition. Only one-person coalitions (2a), two-person and one-person coalitions (2b) and three and one-person coalitions (2c).

4.3.3 Running the Simulation

We distinguish three elements that jointly make up a simulation scenario. During an *iteration*, agents perform several subsequent steps or actions. These steps or actions occur in the iteration's *phases.* Often, one iteration serves as input for the next iteration, to accomplish agent reinforcement learning. Several iterations make up a *simulation run*. Several simulation runs, often each with particular parameter settings, make up a simulation *scenario*. A simulation may, but need not, consist of several scenarios.

To run an iteration, it needs to be set up first. Every iteration starts with an initialisation phase, often followed by a number of phases in which agents interact. Every phase, a number of actions is performed by the agents and the agent environment. Klusch and Gerber (2001) provide a four-phase approach to agent coalition formation during an iteration (note how, somewhat confusingly perhaps, the term 'simulation' here denotes a specific phase in an iteration):

1. **Initialisation:** variables are set to their initial values
2. **Simulation:** simulate possible coalitions and their prospective value
3. **Negotiation:** settle an agreement on the division of payoff
4. **Evaluation:** evaluate agents' ranking. Go back to step 2.

Our simulation scenario follows a similar procedure. Figure 4.3 shows the steps to be taken during each of the four phases Klusch and Gerber identified:

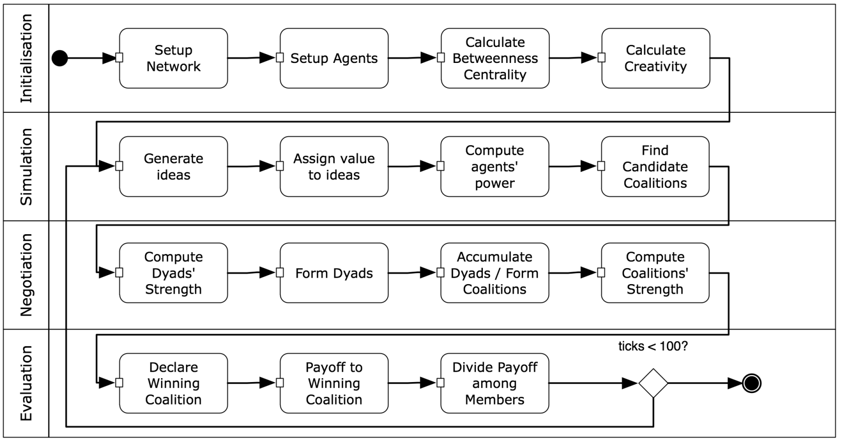


Figure 4.3: Steps to be taken during each of the phases in the simulation.

During the initialisation phase, the network is set up. That is, a network type is chosen and relationships are drawn between agents according to this type of network. Next, agent characteristics (age, personality, etc.) are set to initial values and betweenness centrality and creativity are calculated for each of the agents. Betweenness centrality is calculated using an implementation of the pseudo-code provided by Ulrik Brandes (1994).

$$Cr_i = w3 * Cb_i \tag{1}$$

Where the creativity for agent i, Cr_i, is computed by multiplying the betweenness centrality Cb_i with a predefined weight, $w3$.

The simulation phase comprises several actions to be performed. First, agents generate new ideas. These ideas are given a value, based on the creativity of an agent. We use the following formula to do so:

$$v_{ij} = random(100) + Cr_i \tag{2}$$

Where the value v for idea j of agent i, v_{ij}, is computed by drawing at random a value between 0 and 100 for an idea, and adding the creativity for agent i, Cr_i, to it. We choose to assign a random value to an idea, as we are convinced that anyone can generate a good idea. Other factors may influence the implementation of that idea, but this does not mean an individual cannot generate good ideas, whatever position their position in the organisation. An additional advantage of a random idea value is that it yields dynamics as a result of unpredictable behaviour in simulation of the model.

An agent's power is computed by combining an agent's betweenness centrality, perceived idea value and the actual power of the agent, multiplied by their respective weights. The formula is as follows:

$$P_i(t+1) = w1 * Cb_i + w2 * v_{ij} + w4 * age_i + P_i(t) \quad (3)$$

After updating the power of the agents, the values are normalised, such that every agent has a power value between 0 and 100. At the start of the simulation, $t = 0$, the agent's power is set to a random value between 0 and 100.

Next, each agent computes the scores that other agents have. Similarity to another agent, the power of that agent and the betweenness centrality determine the score of that agent. Similarity is calculated by the following formula:

$$Sim_{ik} = w9 * SimBel_{ik} + w10 * SimGen_{ik} + w5 * SimAge_{ik} \quad (4)$$

Where the similarity in personality between agents i and k, $SimBel_{ik}$, is determined by comparing their Belbin team role. If it is similar, $SimBel_{ik}$ is set to 100. The similarity in gender is computed by looking at the gender of both agents. If they are similar, $SimGen_{ik}$ is set to 100. As the maximum difference in age can be 50, we multiply the age difference between two agents ($SimAge_{ik}$) by 2, in order to have all three similarity measures carry equal weights.

The agent score is calculated by the following formula:

$$Score_j = w8 * Sim_{ik} + w6 * P_i \quad (5)$$

In this case, agent k computes the agent score for each of the other agents. Next, candidate coalitions are looked for, that is, agents that are 'known' through the connections that were set up during the initialisation phase. An agent knows another agent if they are directly connected to each other.

During the negotiation phase, the Shapley value provides a recommendation of candidate dyads. Dyads' Shapley value is computed by summing up the agent scores of the two agents that could form a dyad, minus the strength of the individual agents. The agent chooses to form a dyad with the candidate that is rated highest by the Shapley value.

Subsequently, any two dyads sharing an agent are put into one coalition. As a consequence, all agents that are connected to each other through these dyad connections are put into one coalition. For instance, if agent A and B form a dyad, and agent B and C form a dyad, they together form a coalition that contains agent A, B and C. The coalition's strength is calculated by aggregating the scores of the members of the coalition.

Finally, a winning coalition is declared during the evaluation phase. It is comprised of agents with the highest accumulated strength. Next, the payoff is rewarded to the winning coalition and equally divided among the coalition's members. The individual payoff is then used to update the agent's power. Each agent receives a share of the payoff equal to its share in the coalition's total strength. At this juncture, the current iteration ends. If less than 100 iterations have run, the run returns to the simulation phase; if 100 iterations have run, the simulation run ends.

In the simulation, dynamic behaviour is achieved in two ways. First, the agents generate ideas with a random value. This, in turn, affects the power of an agent. Second, agents that belong to a winning coalition receive a positive update of their power. One may call the result *reputation*.

4.3.4 Parameter settings

We used the following parameters for simulation (Table 4.1):

Table 4.1. Settings for the simulation parameters.

parameter	**setting**
w1	0.45
w2	0.45
w3	0.67
w4	0.1
w5	1
w6	1
w7	1
w8	0.25
w9	0.25
# agents	30
network type	random
network density	0.04
payoff	100
# of runs	100

The values for the weights $w1 - w9$ were found in the literature that we used for the development of our model.

4.4 Results

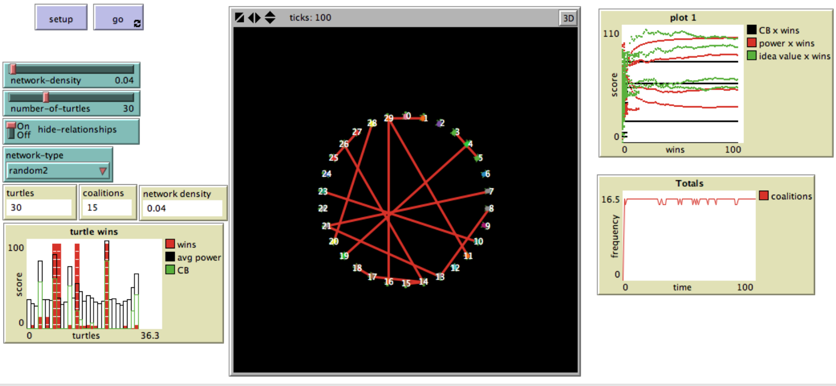


Figure 4.4. Results of the simulation.

Figure 4.4 presents the results of the simulation. Note that the simulation is run in the middle window. Agents that are interconnected by the red lines form a coalition. Same colours for the agents denote that they are in the same coalition.

The histogram entitled 'turtle wins' shows the number of times turtles have won, as compared to their respective betweenness centrality and their average power. Agents are represented on the x-axis 'turtles', starting from the left with agent 0. Red bars indicate the number of wins, black bars indicate the average power per agent, and the green bars indicate the betweenness centrality per agent.

The diagram entitled 'plot 1' shows a number of things. First, the black dots (that show up as a line) indicate the betweenness centrality as a function of the number of wins. The betweenness centrality is stable, as there are no new relationships formed over time. Second, the red dots indicate the power compared to the number of wins. Third, the green dots indicate the idea value compared to the number of wins.

The diagram entitled 'Totals' shows the number of coalitions formed while simulating. As one can see, the number of coalitions has an average of 15.

4.5 Discussion

The results may suggest that there is no direct indicator for a winning agent. Agents with a high score win often and agents with a low score win often. Though, something interesting occurs. If we take a close look at the red dots in plot 1, that is, the number of wins, we see that four agents win all iterations. If we compare this to the histogram 'turtle wins' we see these same four agents represented. The histogram is in the right order of agent number, so if we count from left to right, we see that agent 7, 8, 13 and 21 are winning agents. This is because they are in the same coalition, which is shown in the graphical representation in the middle. What does this mean? It means that their coalition was the strongest one. What made them form a coalition? The Shapley value that recommended valuable peers. This immediately explains why the low-power agents did win during the simulation. They connected to the right agents in their network.

We are well aware that the results obtained with our model and simulation do not necessarily fully apply to reality. First, it is said that the simple simulation models often outperform the more complex ones, as complex models often distort the representation of reality. There are a few things that need to be pointed out, however. Game theory presumes rational play, or rational behaviour among players of the game. Rational play means making optimal decisions, given the actions of other players. Such optimal decisions may maximise the individual or group outcome of playing a game. In reality, players often do not play rationally. Examples include the one-shot version of the Prisoner's Dilemma, in which players are very likely to defect, as they meet only once. Thus, to meet with such irrationalities, we need to adapt the utility mechanism that was used in this simulation. On the other hand, Colman, Pulford, and Rose (2008) state that people do perform team reasoning, as opposed to the irrational behaviour that people are often presumed to have.

Second, the Shapley value has some issues. It does not take into account expected contributions to the coalition. The nucleolus (Schmeidler, 1969; Kohlberg, 1971) does take this into account, and during payoff distribution, it tries to minimise the maximum dissatisfaction of participants in a coalition. We plan to implement this in a new model and compare its results to the current simulation. Also, the Shapley value does not take into account costs for coalition formation. From Lloyd Shapley's perspective, this is quite reasonable, as it is very difficult to capture such costs in a single formula that applies to all situations in which coalitions may occur. Therefore, development of a cost mechanism for coalition formation in networked innovation may be a suitable way to improve our model.

It should be added furthermore, that the Shapley value may be computed in two ways. First, the Shapley value may be computed for people that simultaneously make a move. That is, every person makes a decision whether to cooperate at the same time point. This is the approach we used in the current simulation. We think this method is best for evaluation purposes, in which people decide to cooperate, or vote for someone, after ideas have been generated. Second, the Shapley value may be computed for sequential moves. Coalitions gradually develop in size as more and more people join the coalition. At a certain point, it is not profitable anymore to have someone join the coalition. For instance, a coalition may already be a winning majority, implying that someone joining the coalition will result in dividing the payoff among more people than necessary. For networked innovation, this second way of computing the Shapley value may actually be more promising, but further research into it is required.

Third, for ease of computation, we used Belbin team roles to express someone's personality. Personality may be expressed in more detail using personality traits. In this way we gain a better understanding of which factors influence the perception of similarity among people. This brings us to another point of critique, which is the derivation of the model. Although we did study literature extensively, and used correlation scores from literature for the weights in our model, a tailored approach may be more suitable for our model. Therefore, we plan to test this model on a real dataset of networked innovation. Such a dataset ideally includes personal characteristics and alliances measured over time, and may lead to a more profound model of coalitions in networked innovation. As gaining access to an ideal dataset is likely to be very difficult, we have several options at our disposal. First, viewing co-authoring of academic papers as a kind of innovative collaboration, we plan to use an existing co-authorship network to generate recommendations based on the existing network structure. Second, we plan to develop an 'innovation game' that satisfies the model that we presented in this chapter. Particularly, the game will ask participants to provide access to the network data in their LinkedIn accounts. Additional personal information may contribute to an adequate recommendation of valuable peers for innovation.

Finally, our simulation covered only one scenario with a fixed set of parameter values. Future research should look into the sensitivity of the model results with respect to changes in parameter values. This way the robustness of the results obtained can be assessed. Also, a run consisted of a number of sequential iterations, that is, iterations that adopt the values of a previous iteration as its input (until 100 iterations were run). This however does not show possible variations in the dynamic behaviour of the system. Such variations are to be expected as an agent's creativity is a stochastic variable (equation 2). To estimate the consistency of the dynamic behaviour in the face of this random element, parallel iterations with the same initial values, will also be run.

4.6 Conclusion

In this chapter, we used the Shapley value to generate recommendations of valuable peers in a social network simulation. The algorithm proves to be successful for both low and high scoring agents. Low scoring agents form a coalition with higher scoring agents, thereby loafing on the higher scoring agent's power. By doing so, the higher scoring agents gain a necessary majority for winning the iteration. Thus, both low and high scoring agents profit from the recommendation of valuable peers. The Shapley value, though, presumes rational behaviour of players, which is not always the case. Further research with the present system and improvements of it are suggested.

CHAPTER 5

If We Work Together, I Will Have Greater Power: Coalitions in Networked Innovation

Simulations are especially useful to determine beforehand how certain factors play a role in real life interventions. One can see how the factors affect each other, and how they interact with objects or people by simulating their behaviour. At the NASA space agency, a multi-agent simulation environment was, for instance, used to simulate collaboration and work practice onboard a space station (Acquisti, Sierhuis, Clancey, & Bradshaw, 2002).

This chapter investigates how factors in cooperation networks influence each other, and how sensitive the model is to fluctuations of the variables. It could for instance be that the model can easily be destabilised: a minor change in one variable could have a major effect on the model's resulting behaviour. We implemented a simulation in a multi-agent environment to see how fluctuating variables would affect the dynamics of the simulation. In doing so, we used varied settings for the simulation's factors (*parameter sweeping*) within a specific, predefined range, resulting in 1450 distinct simulations.

This chapter is based on: Sie R. L. L., Bitter-Rijpkema, M., & Sloep, P. B. (submitted). If We Work Together, I Will Have Greater Power: Coalitions in Networked Innovation.

Abstract

The present chapter uses agent-based social simulation to study rational behaviour in networked innovation. A simulation model that includes network characteristics and network participant's characteristics is run using parameter sweeping, yielding 1450 simulation cases. The notion of coalitions was used to denote partnerships in networked innovation. Coalitions compete against each other and several variables were observed for winning coalitions. Close analysis of the variations and their influence on the average power per winning coalition was analysed using stepwise multiple regression analysis. The analysis brought forward two main conclusions. First, average betweenness centrality per winning coalition negatively influences the average power per winning coalition. This implies that having high betweenness centrality as a network participant makes it easier to build a successful coalition, as a coalition needs lower average power to succeed. Second, the number of network participants negatively influences the average power per winning coalition. This implies that in a larger network, it may be easier to form a successful coalition. The results form the basis for the development of a utility-based recommendation system that helps people choose optimal partners in an innovation network.

5.1 Introduction

The rise of the Internet has sparked off a snowballing development of new technologies. In such a rapidly changing world, it is very hard for companies to remain innovative. Only few companies can retain their market share by relying on their internal R&D departments. An increasing number of companies connect to other parties outside the firm to come up with innovations more easily, faster and more cheaply; this is referred to as *networked innovation*. By sharing their knowledge in their social network, they can profit in a number of ways. To illustrate, Google shares its Android mobile platform technology under an open source license. By doing so, others can advance Google's knowledge. Google is well aware that they do not have to invent new technology themselves in order to make money from it. Instead, they use the expert knowledge that is present among the Android developer community and profit from increased adoption and popularity of their Android platform. If good initiatives arise, Google adopts the technology behind it, works together with its originators, or acquires the technology. They fend off risks of financial failure by making effective and efficient use of the knowledge that is present in their network.

The value of networked innovation is emphasised by Cassiman and Veugelers (2006), who found that supportive expertise present in an R&D's social network can boost new product development. Furthermore, Kratzer and Lettl (2008) concluded that people that are on the edge of two social networks have more information, as a result thereof being more creative than others in their network. Ronald Burt

(2004) coined the term *brokerage* for such situations. Perry-Smith (2006) points out the significance of a central network position and weak ties outside the firm to be more creative.

In sum, we can be more creative by profiting from knowledge within our network. Yet, the innovative process does not merely consist of one's creative utterances. Good ideas are often generated, but are for some reason not implemented. Klein and Sorra (1996) point out the importance of skilfulness and commitment for the implementation of innovation. Kotter (1996) suggests a powerful guiding coalition to lead organisational change. Such a coalition is not driven by mere organisational hierarchy, but rather by status, information, expertise, reputations and relationships. The guiding coalition can persuade others in the network to support innovation implementation, which is one of the crucial steps in innovation management (Adamides & Karacapilidis, 2006). A coalition implies a shared intention (commitment) from distinct parties (Ensminger & Surry, 2008; Sie, Bitter-Rijpkema, & Sloep, 2010a). It is necessary to have commitment of all members in order to effectively persuade others in the network. Therefore, we argue that a coalition must have added value for all coalition members as compared to no cooperation (superaddivity). To aid the decision on whom to form a coalition with, we zoom in on the connections that people make during open networked innovation. Forming the right coalitions leads to more innovative power for organisations.

A number of problems arise when in search of coalitions. Firstly, people are not aware of the value of peers in their network neighbourhood (Beham, Kump, Ley, & Lindstaedt, 2010). Secondly, the number of weak ties increases as a social network grows, thereby leading to information overload (De Choudhury et al., 2008). Finally, people lack the cognitive abilities (bounded rationality (Selten, 1998; Simon, 1982, 1991)) to adequately make a choice whom to connect with in order to receive support in adopting their innovation.

In the work presented here, we adopt an agent-based simulation methodology to study coalition formation under rational play in networked innovation. We explicitly limit ourselves to rational play, because the agents' cooperation mechanism is based on game theory. More specifically, prospective connections between agents are viewed as coalitions, and the Shapley value (Hart, 1987; Shapley, 1953) is used to compute the added value of cooperation (forming a coalition) over non-cooperation. Agents exhibit rational behaviour by forming valuable coalitions. The agent-based simulation of networked innovation presented in this chapter allows us to analyse the dynamics of coalition formation in networked innovation. The analysis will lead to a model that helps us predict the behaviour of innovators and its outcomes in a network of innovators. Subsequently, this will result in a recommendation of coalitions in real-life by means of innovation-intervening computer software.

Gilbert, Pyka and Ahrweiler (2001) previously developed a simulation of innovation networks. Their simulation was characterized by: 1) actors, 2) *kenes,* and 3) research strategies. The actors in the simulation represented firms. These firms possessed knowledge and skills, represented by so-called *kenes*. Research strategies dominated the behaviour of the agents and the interaction between agents. That is, an agent could do research and generate knowledge on its own, but it could also form alliances with other agents in order to 'lurk' (copy knowledge and skills) from those agents. Moreover, agents cooperated to generate new knowledge.

We argue that the dynamics of coalitions in networked innovation is very much dependent on the network characteristics and the characteristics of the network's members. By network characteristics we mean the network size and network density (Harary, Norman, & Cartwright, 1965). By the characteristics of the network's members, we mean their age, gender, personality, betweenness centrality and power (reputation). Consequently, the purpose of the present study is to determine whether these have an influence on the power and successfulness of coalitions. A detailed description of the method of simulation and our model will be presented in the next section. Thereafter, we provide the results of our simulation. Next, we analyse the results using stepwise multiple linear regression, and we will discuss these results in the subsequent section. We conclude with some final thoughts and suggestions for future work.

5.2 Methods

5.2.1 Simulation scenario, iterations and phases

We run our simulation using the Netlogo simulation environment. It provides a means to do agent-based social simulation. Agent-based social simulation is an application of two areas, namely agent-based computing and computer simulation to a third area, social science (Davidsson, 2002). Agent-based computing is mainly aimed at the interaction between distinct computer software programs called *agents*. The agents can represent for instance computer systems in NASA space missions (Clancey, Sierhuis, Kaskiris, & Van Hoof, 2003; Seah, Sierhuis, & Clancey, 2005). Events within the (space) environment can be picked up by the agents and acted upon. Computer simulation is a method by which computers can simulate real world behaviour. Unlike agent-based computing, computer simulation does not necessarily employ agents. It uses, for instance, statistical models and Bayesian models to simulate and study the behaviour of liquids (Allen & Tildesley, 1999). Agent-based social simulation allows one to study the dynamics of social interaction such as networked innovation, without the need to implement an intervention system in practice to pilot its workings. This is especially useful if researchers have a one-shot chance of intervening, when intervention is very costly, or when experimental participants are scarce.

The agent-based social simulation that we developed comprises a *simulation scenario*. A simulation scenario is a workflow, or a number of actions that has to be performed during the simulations. Actions can be performed multiple times, and they often take place in pre-defined sequences. When multiple sequences are run in a simulation, we call them *iterations.* An iteration often influences the subsequent iteration by means of *reinforcement*, as is the case with our simulation. An iteration consists of multiple *phases*, to distinguish different types of activities performed during the iteration. During an iteration, we start off with an initialisation phase to set up the agent's and environment's parameters; this is followed by a number of phases in which the agents interact. Akin to a simulation of agent coalition formation by Klusch and Gerber (2002), we distinguish four phases (as depicted in Figure 5.1):

1. Initialisation: The agent and environment parameters are set up
2. Simulation: The candidate coalitions are determined
3. Negotiation: Coalitions are formed
4. Evaluation: The winning coalition and reinforcement is determined

5.2.2 Initialisation

The simulation commences with setting up the network of agents given a predefined network density. Also, the nodes within the network represent individuals and the edges form their relationships. Two individuals are said to be related when the agents are known to each other. Based on their position in the network, the agents' betweenness centrality (Brandes, 1994) is estimated. Betweenness centrality tells us how dependent others are on an individual in a network. For instance, when we have two companies A and B, and only one person in company A connects to company B, then the employees in companies A and B are very much dependent on that single person in terms of information exchange. As a result, that person will have high betweenness centrality. Intuitively, having such a good network position leads to increased power. Also, high betweenness centrality will increase the creativity of an agent.

5.2.3 Simulation

During the simulation phase, the initial parameters and the calculations of betweenness centrality and creativity will be used to let the agents generate new 'ideas'. The ideas are abstract and do not own any content. They receive a value based on the creativity calculation performed in the initialisation phase. Based on the idea value and the betweenness centrality, an agent's power is determined. An agent that has high power is more likely to convince others of the value of an idea. Besides, if it has high betweenness centrality, it may have more decision power, as other agents are dependent on this agent. Power and social similarity (age, gender, personality) (Ibarra, 1992; McPherson et al., 2001) contribute to the likelihood that an agent will be selected for cooperation, the so-called agent score. For instance, if agent A has high power and is very similar to agent B, then agent B will most likely choose agent A to cooperate with (and form a coalition).

5.2.4 Negotiation

We use the Shapley value, a measure well known in game theory, to calculate the value of prospective coalitions. The Shapley value calculates the added value of forming a coalition with another agent over going at it alone. It must be noted that a coalition must be at least as strong as the accumulated strength of its members (superadditivity). In fact, a coalition must be stronger than the accumulated strength of its members (monotonicity). The latter reflects that in real life one inherently needs support to have one's idea accepted by the community. To do so, we form coalitions (Kotter, 1996). As opposed to humans, agents always play rationally, and thus choose to form a coalition with the highest-scoring prospective coalition.

5.2.5 Evaluation

Finally, a winning coalition is determined, that is, the coalition that has the highest accumulated power. Payoff in the form of additional power (in the next iteration) is given to the agents of the winning coalition. This gives us insight into the overall emergent behaviour in networked innovation. More specifically, we see how agent power changes, and how this influences the formation of coalitions and the structure of coalitions. In sum, the simulation expresses dynamic behaviour in two ways. First, the agents generate ideas based on their creativity, plus a random value. In turn, this affects the power of an agent. Second, agents that belong to a winning coalition receive a positive update of their power. One may call the result reputation.

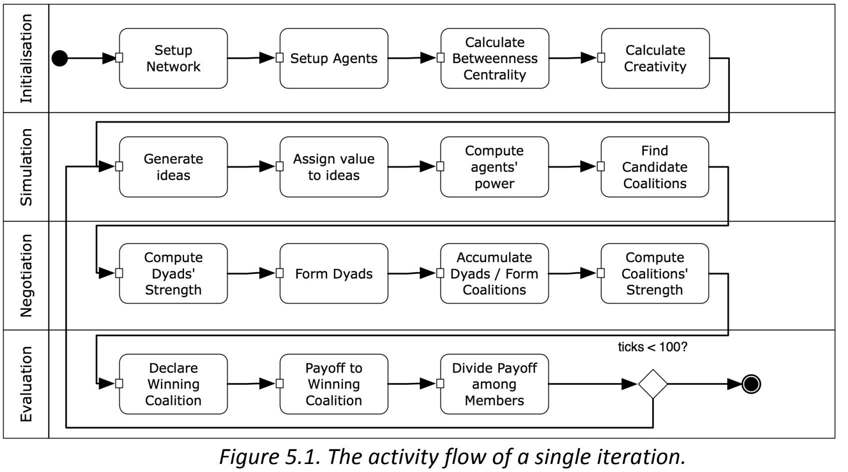


Figure 5.1. The activity flow of a single iteration.

5.2.6 Simulation model

The above overview of iterations and phases does not by itself make a simulation run. In agent-based simulation, agents have an internal reasoning model. This model may be regarded as the internal reasoning structure of an agent and allows an agent to perceive other agents and its environment. Figure 5.2 shows the internal reasoning structure of our agents. Note that every agent is the same by nature, but initial parameters such as gender, age and personality may vary per agent.

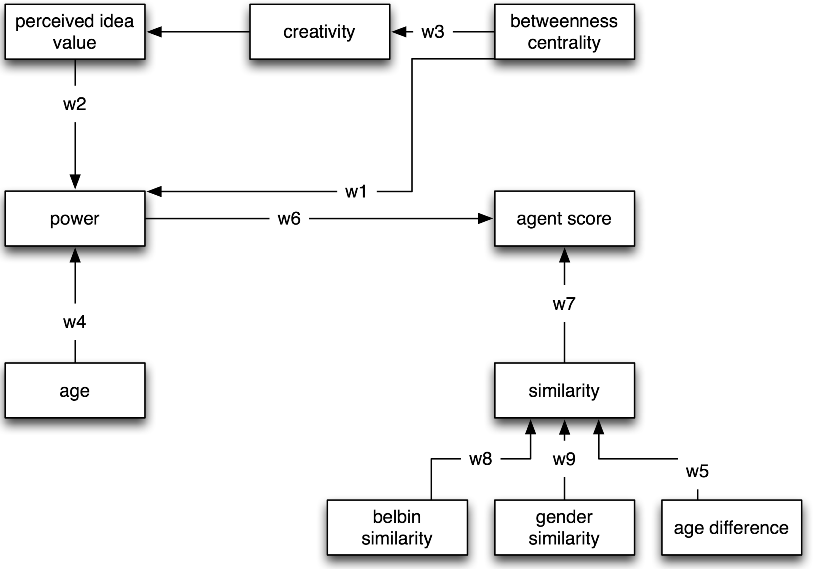


Figure 5.2: The simulation model; for a detailed description, see text.

5.2.7 Weights

There are two factors that mainly influence the decision to form a coalition: 1) power and 2) homophily. Power and the similarity between two individuals (homophily) directly influence the agent's score. The agent's score represents the likelihood that agent A is interested in forming a coalition with agent B. There are seven other factors that indirectly contribute to an agent's score through the two central factors. The factors (including the agent score) are connected through weights, to indicate the effect of one factor on another. The value of the weights is not decided upon arbitrarily; literature was used to determine their value. The value per weight may vary, as is shown in Table 5.1. Note that it is not a goal to perfectly and precisely display reality in this model. To do so, we would have to include all possible factors and the exact weights between them to exhibit the appropriate behaviour. We merely seek to simulate behaviour that sufficiently closely resembles reality. In fact, it is common knowledge among agent-based modelling researchers that a more complex model often results in a less representative simulation of a situation. In our practice, this means we included relatively few factors in our simulation model to maximise outcome.

Table 5.1. Weights, their values, and origin in literature.

Weight	Value	Literature
w1	0.45	(Brass, 1984; Ibarra, 1992, 1993a; Krackhardt, 1990a; Perry-Smith, 2006; Simon, 1982)
w2	0.45	(Klein & Sorra, 1996)
w3	0.67	(Kraatz, 1998; Kratzer & Lettl, 2008; Tsai & Ghoshal, 1998)
w4	0.1	(Burkhardt & Brass, 1990)
w5	1	(Ibarra, 1993a; McPherson et al., 2001)
w6	1	(Ibarra, 1992; Kotter, 1996)
w7	1	(Ibarra, 1993a; McPherson et al., 2001)
w8	0.25	(Ibarra, 1993a; McPherson et al., 2001)
w9	0.25	(Ibarra, 1993a; McPherson et al., 2001)

The concept of betweenness centrality originates from Social Network Analysis (Wasserman & Faust, 1994) and is used to express someone's position in a network. It measures how dependent others are on a target node (individual) in a network. It is computed by the number of shortest paths between individuals that pass through a node, as a proportion of all shortest paths possible. In our case, betweenness centrality measures how dependent people are on one another if they want to connect. People cannot form a coalition if there is no path that connects them. If an agent possesses high betweenness centrality, agents very likely have to pass it to reach any one agent in the network. Betweenness centrality has an impact on a number of factors. First, people that are on the edge of two networks, and thus have higher betweenness centrality, are more likely to be creative or innovative than others (Kratzer & Lettl, 2008; Tsai & Ghoshal, 1998). To take this one step further, interorganisational ties may advance social learning, thereby contributing to organisational growth (Kraatz, 1998). Secondly, central individuals are found to be more powerful (Brass, 1984; Ibarra, 1992, 1993a; Krackhardt, 1990; Perry-Smith, 2006; Simon, 1982).

Age and perceived value of an idea also influence power. Age is found to correlate positively with power (Burkhardt & Brass, 1990). Klein and Sorra (1996) suggest that 'innovation-values fit', the extent to which an innovation (idea) fits the perceiver's values, influences support for an innovation. In our model this is represented by the perceived value of an idea.

Homophily, the similarity between people, has a positive influence on support and friendship relationships (Ibarra, 1992). Various types of homophily may exist, such as age and gender (McPherson et al., 2001). For our model, we use age, gender and personality to express similarity. Besides, a change in thought must be led by a group that has decision power and persuasive power. Kotter (1996) denotes such a group by a *guiding coalition*.

5.2.8 Variables

Age is represented as a random value between 15 and 65, the so-called 'working age' of people. Gender is represented as a random value of 0 (female) or 1 (male).

Personality is difficult to represent. Multi-attribute personality scores such as the Big Five personality traits have been considered, but for the time being, we choose to use the Belbin Team Roles (Belbin & Belbin, 1996). The nine Belbin profiles express the role of a person within a team. Use of these predefined team roles eases the computation of similarity.

Agents have a power attribute, which corresponds to their power in the model. Agents' ultimate score is influenced by both their power and their similarity to other agents.

Table 5.2. An overview of the variables, their initial value, value range, and how they increment.

Variable	Variable abbreviation	Range	Increment	Initial value
Betweenness centrality	Cb_i	1 – ∞	n/a	n/a
Creativity	Cr_i	0 – 100	progressive	n/a
Power	P_i	0 – 100	progressive	
Gender	Gen_i	0 = female, 1 = male	n/a	random
Age	Age_i	15 – 65	1	15 + Random(50)
Belbin personality	Bel_i	1 – 9	1	Random(9)
Perceived idea value	v_{ij}	0 – 100	progressive	n/a
Similarity	Sim_{ik}	-50 – 50	1	n/a
Belbin similarity	$SimBel_{ik}$	0 – 100	100 (Boolean)	n/a
Age similarity	$SimAge_{ik}$	0 – 100	1	n/a
Gender similarity	$SimGen_{ik}$	0 – 100	100 (Boolean)	n/a

5.2.9 Formulas

Some of the variables in Table 5.2 do not have an initial value. They are calculated during the simulation. Their respective formulas are shown in Table 5.3.

Table 5.3. Formulas used for determining intermediate value and weights.

#	Name	Abbreviation	Formula	Variables
1	Creativity	Cr_i	$Cr_i = w3 * Cb_i$	w3, Cb_i
2	Idea value	v_{ij}	v_{ij} = random(100) + Cr_i	Cr_i
3	Power (update)	$P_i(t+1)$	$P_i(t+1) = w1 * Cb_i + w2 * v_{ij} + w4 * age_i + P_i(t)$	w1, Cb_i, w2, v_{ij}, w4, age_i, $P_i(t)$
4	Similarity	Sim_{ik}	$Sim_{ik} = w9 * SimBel_{ik} + w10 * SimGen_{ik} + w5 * SimAge_{ik}$	w9, $SimBel_{ik}$, w10, $SimGen_{ik}$, w5, $SimAge_{ik}$
5	Agent score	$Score_j$	$Score_j = w8 * Sim_{ik} + w6 * P_i$	w8, Sim_{ik}, w6, P_i

5.2.10 Procedure and data collection

During execution of the simulation model we set two parameters using *parameter sweeping* to see how they influence coalition formation among agents: 1) *network density* (number of relationships divided by the number of total possible relationships) and 2) *number of turtles* (number of network participants). In parameter sweeping, we vary the values for these independent variables in a structured way within a predefined range. Parameter sweeping allows one to report and analyse the dynamics of simulations within a wide parameter space. It requires little human effort, as one does not have to enter all parameter combinations manually (Brueckner & Van Dyke Parunak, 2003). The range of the *network density* parameter varies from .01 to .0.05 with an increment of .01 (5 values). The range of the *number of turtles* parameter varies from 2 to 30, with an increment of 1 (29 values). This results in 145 possible combinations of parameters. Each combination of the parameters (simulation run) is executed 10 times to yield stable results. This implies that in total we run 1450 simulations. We observe the following parameters for their fluctuations and to find relationships with the average power per winning coalition:

- **network density:** The extent to which relationships are formed as a function of all possible relationships
- **number of turtles:** The total number of participants in the network
- **average-betweenness-per-winning-coalition:** We measure the average betweenness centrality of the members of a winning coalition to see if there is a relationship between the independent variables and this dependent variable
- **average-idea-value-per-winning-coalition:** We measure the average idea value of the members of a winning coalition to see if there is a relationship between the independent variables and this dependent variable
- **max-power-per-winning-coalition:** We measure the highest power of a member of a winning coalition to see if there is a relationship between the independent variables and this dependent variable
- **max-idea-value-per-winning-coalition:** We measure the highest idea value of a member of a winning coalition to see if there is a

relationship between the independent variables and this dependent variable

5.2.11 Data Analysis

We will analyse the simulation results in two steps. First, we use multiple regression analysis to create a model that predicts the influence of independent variables on the dependent variable *average power per winning coalition*. Second, we investigate the validity of the model by analysing the correlation between its residuals (Durbin-Watson statistic), as regression assumes absence of such correlation. A Durbin-Watson statistic near 2 implies that there is no correlation between adjacent residuals. When using regression, it is key that the residuals be independent.

5.3 Results

A total of nine variables were exported from the simulation to determine if and to what extent they predicted the *average power per winning coalition*. The correlation coefficients for the variables using Pearson Bi-variate correlation are provided in Table 5.4. High correlation exists between the pairs {*total number of coalitions*, *number of turtles*}, {*max betweenness per winning coalition*, *average betweenness per winning coalition*}, {*max idea value per winning coalition, average idea value per winning coalition*}. Moderate correlation exists between the pairs {*max betweenness per winning coalition, average power per winning coalition}.*

Table 5.4. Correlation coefficients for each of the variables.

	average power per winning coalition	network density	number of turtles	average betweenness per winning coalition	average idea value per winning coalition	max power per winning coalition	max idea value per winning coalition
average power per winning coalition	1.00						
network density	-.28	1.00					
number of turtles	-.59	.00	1.00				
average betweenness per winning coalition	-.57	.33	.41	1.00			
average idea value per winning coalition	.05	.07	.14	.29	1.00		
max power per winning coalition	.26	.12	-.08	.11	.30	1.00	
max idea value per winning coalition	-.38	.22	.41	.56	.76	.29	1.00

The outcome of multiple regression analysis using the stepwise method is presented in Table 5.5. Table 5.5 shows the predictive values for the variables of the best scoring model in which six variables were included.

Table 5.5. Multiple regression analysis of the simulation for average power per winning coalition. Six variables were included in the model, sorted in the order they were entered.

	b	**SE b**	**ß**
Constant	42.42	2.95	
Number of turtles	-.44	.03	-.31*
Average betweenness per winning coalition	-.33	.02	-.27*
Max power per winning coalition	.56	.03	-.29*
Network density	-115.39	13.43	-.14*
Average idea value per winning coalition	.31	.02	.50*
Max idea value per winning coalition	-.24	.01	-.54*

Note. R^2= .68. * p < .001

Using the stepwise method, a significant model emerged (*$F_{6,1443}$ = 514,675, p < 0.001).* As shown in Table 5.5, two variables have slightly larger influence on the *average power per winning coalition*: *number of turtles* and *max betweenness per winning coalition.* The R^2 shows that the variables account for 68% of the predictability of *average power per winning coalition*. The variable *network density* yielded no significant results. To make sure no auto-correlation exists we used the Durbin-Watson statistic. A Durbin-Watson value of 1.80 (near 2) implies that there is no auto-correlation.

5.4 Discussion

The correlation scores in Table 5.4 inform us about the co-occurrence of variables. We see that, as the network size (number of turtles) increases, so does the total number of coalitions. This is to be expected, as a larger network implies more candidate connections between people. However, a decreasing network density may have a counter effect on the number of coalitions that is formed. Most important for the multiple regression analysis is that there is no relationship between the independent variables (predictors) number of turtles and network density. Otherwise, the multiple regression model could not be written in the form of $Y = c+b_1X_1+b_2X_2$.

The R^2 of .68 indicates that the variables in Table 5.5 account for 68% of the predictive value of the *average power per winning coalition*. Our results are in contrast with literature that shows that betweenness centrality influences power within networks (Brass, 1984). Table 5.5 shows that the average betweenness centrality of a winning coalition has a negative influence on the average power of a winning coalition. The study by Brass, though, was not designed to take into account innovation within networks, a special case of social networks. Subsequently, we see a positive influence of the *average idea value per winning coalition* on the power of a coalition, in line with our reasoning.

Another value that stands out is the network density. The reason for this is that we used relatively small variations of the network density, thus compensating for the supposedly high influence observed in Table 5.5.

A notable observation we find in a combination of Tables 5.4 and 5.5. *Average betweenness per winning coalition* correlates moderately high with the *average power per winning coalition* (-.57). Besides, it negatively influences the *average power per winning coalition*. A high betweenness often means that one has a lot of contacts in one's social network that others do not have. Having lots of contacts implies one cannot maintain close relationship with all contacts, leading to an increased number of weak ties. Literature is suggestive of the strength of weak ties (Granovetter, 1973; Hauser, Tappeiner, & Walde, 2007) in social networks (Granovetter, 1973). Especially, networked learning (Jones, Ferreday, & Hodgson, 2008) and networked innovation (Burt, 2004; Hauser et al., 2007) value weak ties as predictors of successful cooperation in networks. Our results imply practically the same; Table 5.5 shows that average betweenness per winning coalition negatively influences the average power per winning coalition. In other words, having high betweenness centrality makes it easier to build a successful coalition as one needs a lower average power to succeed.

Another interesting observation lies in the negative influence of the *number of turtles* on the *average power per winning coalition* (Table 5.5). This implies that as the network size increases, it becomes easier to build a successful coalition. Although other factors may influence the process as well, we may conclude that it may be easier to form a successful coalition in a larger network.

There are two implementations of the Shapley value. First, we have the situation in which all agents form a coalition at once, the one that we used in this simulation. Second, the agents may join a coalition one after another. In case of a high-betweenness agent attracting a lot of partners, we could consider using the second method of coalition formation to further optimise the simulation. Besides improving the way the Shapley value is calculated and used for the formation of coalitions, we may decide to implement the nucleolus. The Shapley value does not consider the expected contribution of an agent to a coalition, whereas the nucleolus (Schmeidler, 1969) does. During payoff distribution, the nucleolus tries to minimise the maximum dissatisfaction of participants in a coalition.

5.4 Conclusion

The present study investigated whether network characteristics and network member's characteristics influence the average power per winning coalition. To aid people in their search for optimal coalitions, we studied the dynamics of coalitions in networked innovation. We ran a simulation of networked innovation under rational behaviour (to yield optimal decisions), and monitored the variable variations. Multiple regression analysis led to a model that predicts the average power per winning coalition as a function of network size and network density.

The current study allows us to make two interesting observations. First, average betweenness negatively influences the average power per winning coalition. This means that having high betweenness centrality makes it easier to build a successful coalition, as one needs lower average power to succeed as a coalition. Second, the number of network participants negatively influences the average power per winning coalition. This implies that in a larger network, it may be easier to form a successful coalition.

The regression model presented in this chapter offers interesting uses. Our simulation presumes rational play by network participants. In other words, optimal decisions are made concerning the formation of coalitions. Assuming rational play, we compute how coalitions should ideally be formed within networked innovation. An important implication of this model is that we can assist in real life networked innovation by recommendation of optimal coalitions (with a necessary average power or betweenness centrality), given that we know what the network density and network size are.

5.5 Future Work

The model presented in this work was based on extensive literature review. The research articles that we studied employ empirical methods to determine if and what relationships between variables exist. We combined the outcomes of several influential studies to develop a simulation model. We programmed agents on an individual level to study the emergent dynamics of networked innovation (macro level), an approach that is characteristic for agent-based social simulation. The next step in the process of deriving a model that correctly describes reality is the validation of the model. We plan to validate our model by testing its behaviour against empirical data. Subsequently, we will use the model to generate optimal coalitions for innovation in networks in an empirical setting.

CHAPTER 6

To whom and why should I connect? Co-author Recommendation based on Powerful and Similar Peers

This chapter is a first user evaluation of our COalitions in COOperation Networks (COCOON) system. Similar to the simulations in Chapters 4 and 5, new connections are formed between network members, based on the network position and similarity of network members. COCOON aims to help researchers find the right co-author for their next article. To cooperate well, co-authors need to have some sort of similarity, a common ground that unites them and the topics in the article. Also, the cooperation needs to be successful, that is, an article should very likely be accepted by the reviewers.

One way of accomplishing a high chance of acceptance is by including co-author power (authority) in the recommendation algorithm. If we search for a co-author, and we want the article to have a higher chance of acceptance, we should connect to a peer in the network that has authority. The recommendation algorithm combines network authority with interest similarity between candidate co-authors.

This chapter is published as: Sie R. L. L., Drachsler, H., Bitter-Rijpkema, M., & Sloep, P. B. (2012). To whom and why should I connect? Co-author Recommendation based on Powerful and Similar Peers. *International Journal of Technology Enhanced Learning (IJTEL), 4*(1), 121-137, DOI: 10.1504/IJTEL.2012.048314

Abstract

The present chapter offers preliminary outcomes of a user study that investigated the acceptance of a recommender system that suggests future co-authors for scientific article writing. The recommendation approach is twofold: network information (betweenness centrality) and author (keyword) similarity are used to compute the utility of peers in a network of co-authors. Two sets of recommendations were provided to the participants: Set one focused on all candidate authors, including co-authors of a target user to strengthen current bonds and strive for acceptance of a certain research topic. Set two focused on solely new co-authors of a target user to foster creativity, excluding current co-authors. A small-scale evaluation suggests that the utility-based recommendation approach is promising, but to maximize outcome, we need to 1) compensate for researchers' interests that change over time, and 2) account for multi-person co-authored papers.

6.1 Introduction

We often see that creative ideas are lost during the innovation process. Good and creative ideas are generated, but we see a lack of support and commitment of valuable ideas by other employees. We argue that the innovation process is, to a large extent, similar to organisational change processes and can thus profit from insights in this field of research. Both innovation and organisational change aim to alter and optimise the way we think, act, or make things. Furthermore, the contexts of both change processes are recognised by a predominant, common intention and a shared identity (community of practice (Wenger, 1999)). The innovation process tries to advance current state-of-art products, services or technologies, while organisational change aims to improve the current practice.

Both innovation and organisational change suffer from similar problems. One of the main reasons organisational change fails is the lack of a guiding coalition (Kotter, 1996). To successfully change an organisation, it is important that a change be adopted by several powerful employees. Innovation implementation often fails because the innovation does not fit the values of the employees (Klein & Sorra, 1996). Thus, both experience a lack of support and commitment. For example, the Post-It note was not perceived as valuable by the 3M company until the employee that came up with the idea started spreading the notes among secretaries. The secretaries kept asking for more of these notes, which eventually persuaded the Marketing and Strategy department (West, 2002); A guiding coalition was formed by the inventor and the secretaries.

The solution to effective change and innovation implementation seems obvious. We have to find the right, powerful peers to connect to. Please note that by powerful, we do not mean powerful by hierarchy per se. Powerful peers can be think-alikes, for example, people that have the ability to persuade others, or senior employees. Though, a number of problems hinder one from finding the right peers. Firstly, people face an abundance of other people that they can connect to (information overload (De Choudhury et al., 2008)). Secondly, people are *boundedly rational* (Selten, 1998; Simon, 1991); they lack the cognitive abilities to determine the value of candidate cooperating peers, also due to *lack of awareness* (Reinhardt, Mletzko, Drachsler, & Sloep, 2011). Thirdly, people are self-interested (Kau & Rubin, 1979; Ratner & Miller, 2001); they need an incentive for cooperation. In other words, they need to know what the added value is of cooperating with others. Indeed, other people hold complementary knowledge. Therefore, many recommender approaches nowadays focus on recommendation of peers to discover complementary knowledge (Beham, Kump, Ley, & Lindstaedt, 2010; Vassileva, McCalla, & Greer, 2003).

We argue that the above problems result in non-optimal outcomes in research collaboration. In this study, we investigate a co-authorship network in order to recommend possible future cooperative writings. Other studies acknowledge the same problems in research and try to solve them by raising awareness (Reinhardt et al., 2011), designing a platform to mediate collaboration (Ullmann et al., 2010) or recommending scientific events (Klamma, Phnam, & Cao, 2009).

Our approach is inspired by two thoughts: 1) networked innovation and learning and 2) utility theory. With respect to the first thought, we regard cooperative writing of research papers (network interactions) as a joint learning and innovation action. By cooperatively writing a paper, the authors necessarily connect to each other. Together, the authors (nodes) and paper writing (edges) form a network of co-authors.

With respect to the second thought, we use the prospective value (utility) of candidate cooperation to recommend peers. Expected utility calculations originate from game theory. It widely gained popularity when John von Neumann and Oscar Morgenstern published their book Theory of Games and Economic Behaviour back in 1945 (Von Neumann & Morgenstern, 1945). As the title suggests, it was initially used for the analysis and prediction of economic behaviour. Over the last decades, however, other fields of research have applied game theory, including computer science (Abdallah & Lesser, 2004; Jonker, Robu, & Treur, 2007; Klusch & Gerber, 2002; Sie, Bitter-Rijpkema, & Sloep, 2010b). In short, the prospective value of a peer is computed by the network position of a peer, and the similarity to that peer in terms of the keywords that they use.

To this end, we extract metadata from a publication database that uses the DSpace software. DSpace is a publication database in which researchers can upload their publications. Especially for researchers, it is important to reach out beyond the borders of their own university, connect to other researchers, and gain general acceptance through citation of their work. DSpace is based on the Open Archives

Initiative, and offers a predefined, structured method for publishing to, and openly extracting metadata from the database. The database at hand consists of a set of presentations, research papers, and project deliverables. As noted earlier, the authors of the documents form a network of co-authors and keywords that are provided during submission of the document to the database are used to compute similarity between authors in terms of research interest.

Two sets of recommendations will be shown to the participants. Recommendation Set 1 includes people that the target user has written with so far, and recommendation Set 2 excludes these people. The main question we ask ourselves is: *How well do participants perceive a recommendation that is based on keyword similarity and network information to be?*

The outline of this chapter is as follows. In Section 6.2, we discuss the research methodology. We describe the dataset that we apply, the recommendation algorithm and the method of evaluation. Section 6.3 presents the results of our evaluation. In Section 6.4, we discuss the results of the evaluation, and in Section 6.5, we draw our conclusions and provide an outlook for future work.

6.2 Method

6.2.1 Data Collection

The dataset that we use is extracted from a DSpace publication database. The database comprises 1009 research publications, 518 presentations and 357 project deliverables. Every submission is placed in a certain category, that is, the department where it was written. Table 6.1 provides a numerical overview of the database. As for this dataset, some of the departments do not have a long history of research publications. For example, departments A, B and C have been doing research for over ten years, whereas department D was founded in 2008. Department F and G started doing research in 2004. Differences in the amount of data may influence the resulting recommendations.

Table 6.1. Numerical overview of the publication database.

Department	publications	presentations	deliverables
A	373	247	184
B	280	170	131
C	155	10	0
D	62	89	42
E	3	2	n/a
F	102	n/a	n/a
G	13	0	n/a
H	43	n/a	n/a
I	21	1	n/a
Totals	1009	519	357

The following metadata is provided by the author when an individual submission is posted to the database:

- Unique identifier
- Timestamp: date and time of submission
- Creators: the authors
- Descriptions: APA reference, sponsors
- Language
- Title
- Subjects: keywords that specify the contents
- Type: Journal paper, conference paper, book chapter, etc.

Every submission contains one or more authors. By cooperatively writing an article, the authors are inherently interconnected. These connections can be used to form a so-called *one-mode complete network* of co-authors. This is, however, different than the usual citation networks in which citations between articles are used to generate a network. Besides, we can construct other types of networks to enhance our algorithm, such as relationships based on the department the article was written, the type of submission, or the keywords that are used to describe the article. For the present study, we focus on the keywords to measure similarity between authors, but we are planning to further optimise performance by putting the other alternatives to use as well.

The extraction of authors is done as follows. The DSpace software is based on the Open Archives Initiative (OAI) (Lagoze & Van de Sompel, 2001). The OAI provides a protocol for metadata harvesting (OAI-PMH) that can be used to extract submissions from the dspace.ou.nl website. A HTTP request is made to the DSpace's OAI-PMH containing the identifier of a subset (collection) of DSpace. The DSpace OAI-PMH returns an XML file that contains all submissions in that subset of the DSpace website. Next, this XML file is read out by a PHP script that splits every entry (submission) into several types of data that are each stored in separate tables in a MySQL database. This repeated for every collection of submissions in DSpace. The MySQL database model is shown Figure 6.1.

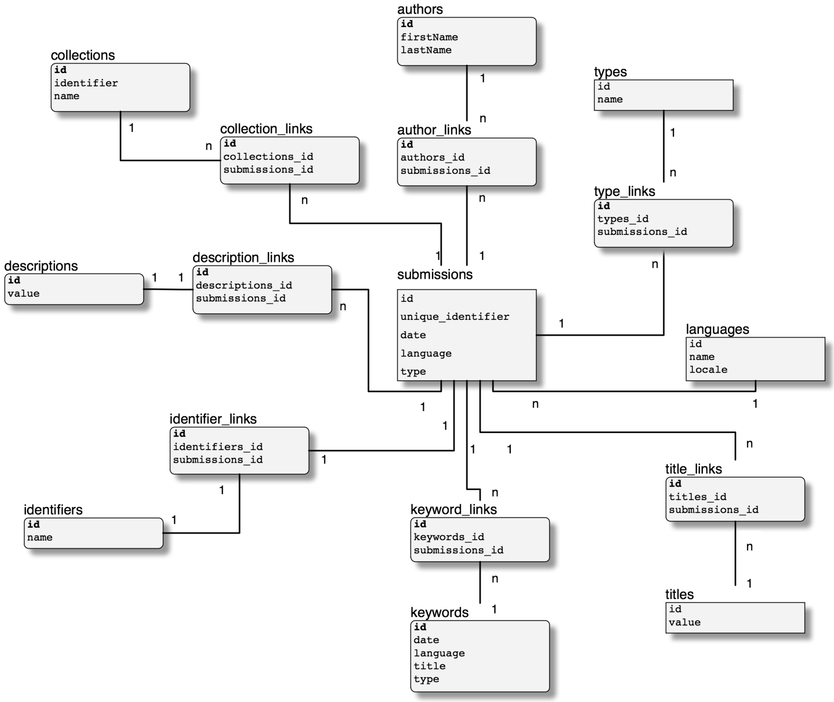


Figure 6.1. MySQL database model for the DSpace data

Figure 6.1 shows that authors and submissions are stored separately. Authors can link (author_links) to multiple submissions, as they store multiple submissions. Submissions can link (author_links) to several authors, as multiple authors can contribute to a single submission. In this way, we can create a co-authorship network by performing the following actions: 1) get an author's submissions by retrieving all author links to submissions, 2) for each submission, look for all author links to authors, 3) save this as a network connection, 4) repeat step 1-3 for every author in the database, while keeping in mind not to process duplicates. A more formal description of this algorithm is shown in Table 6.2.

Table 6.2. Algorithm for extraction of the data.

Algorithm 1: Co-author extraction in an unweighted, bidirectional graph
// make an empty stack of connections between authors
P[v,w] ← empty stack, v,w ∈ V;
foreach submission s ∈ S **do**
foreach author a of s **do**
foreach author b of s do
// if a and b are not equal, and they are not in the stack of connections
if a ≠ b and a,b ∉ P[v,w] **then**
// save the connection to the stack
push a,b →P[v,w];
end
end
end
end

6.2.2 Recommender System

We envisage the workflow of our recommender system as follows:

1. Co-authors are extracted from papers to create a co-author network
2. Authors receive a value, based on their network position, and their similarity to the query author
3. Candidate dyadic[2] connections utility-based value
4. The users receive a ranked list of researchers

Figure 6.2 depicts the recommendation process. Numbers correspond to the above list.

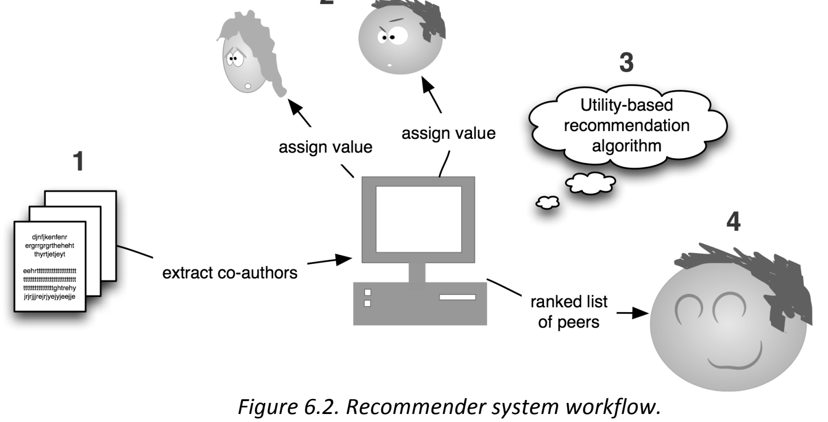


Figure 6.2. Recommender system workflow.

[2] A dyad is another name for two people that belong to the same social group, in this example candidate co-authors.

After data collection in step one of the workflow, in step two the authors receive a value based on the network position of the authors. To be more precise, betweennesss centrality (Brandes, 1994) is used to calculate to what extent other authors are dependent on an author in terms of information flow. In formal terms, betweenness centrality stands for the number of times a node (author) is on the shortest path of any pair of nodes relative to the total number of shortest paths in the network. In case of co-authorship networks, betweenness centrality stands for the extent to which other authors are dependent of a certain author when disseminating research ideas within the network.

Individuals that have high betweenness centrality in the network are found to be more powerful (Ibarra, 1992, 1993; Krackhardt, 1990; Perry-Smith, 2006; Simon, 1982). In a co-authorship network, we can explain this in two ways. First, individuals that are often on the edge of two networks (high betweenness centrality) have more access to new viewpoints. Therefore, they are able to apply knowledge from one domain to another domain, thereby being more creative (Burt, 2004). Second, individuals that are on the edge of two networks have power over the information flow between the two networks. This gives them more status and power (Krackhardt, 1990). This often shows from an individual's hierarchical position in the organisation in relation to their betweenness centrality. Preliminary observation of our dataset shows that individuals that are high in the organisational hierarchy also have a high betweenness centrality. This leads us to believe there is a relation between key job positions and the betweenness centrality of an individual in an organisation. The betweenness is spread like a *long tail* distribution; Few authors have high betweenness, and many authors have low betweenness (Figure 6.3).

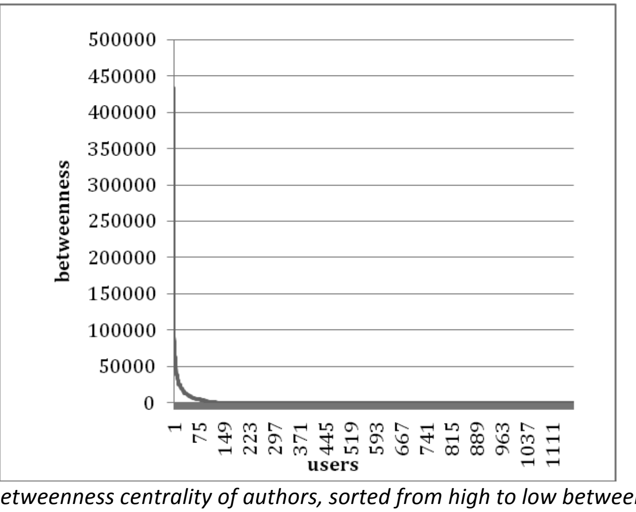


Figure 6.3. Betweenness centrality of authors, sorted from high to low betweenness.

Next, we compute the similarity between authors. High similarity, in gender for instance, is found to be an indicator for good relationships (Ibarra, 1992), and this is supported by research on homophily and friendships (Lazarsfeld & Merton, 1954; McPherson et al., 2001). To measure similarity, we first have to identify individuals within the network. For each of the authors, we look at their submissions and the keywords that they have used in these submissions. We prefer the use the keywords over the title or the contents. Akin to this chapter's title, authors sometimes use appealing sentences to trigger a potential reader's attention. As a result, mapping the title to the interests of the authors may not always work like we want to. Processing the full content of papers often takes too much time, especially when the database size increases, and can therefore not be used to compute real-time recommendations. The keywords that authors use to identify their paper is in our opinion the best way to determine their interest and expertise and compute real-time recommendations.

We use the overlap of expertise (keywords) between individuals to express their similarity. In detail, this is done by retrieving the keywords for every paper an author has written. These keywords per author are then used to compute the term frequency inverse document frequency (TFIDF). That is, each keyword receives a value, but keywords that are used often receive a lower value. For instance, since a large group of people in our dataset work in the field of technology-enhanced learning, the term technology-enhanced learning shows up very often as a keyword in papers. Our recommender system will take this keyword into account, but it receives a lower value. In this way, we can recommend more unique co-authors, rather than recommending one author (that used the keyword technology-enhanced learning very often) to everyone. Afterwards, the vector similarity between authors is computed by treating the set of keywords per author as a vector.

In step three, we use a utility-based algorithm for our recommendation of peers. The algorithm uses the predictive value of a peer in the network to estimate whether or not cooperation should be pursued. This value is estimated using the two types of similarity from step two. The two similarities are different in size, however. For this experiment, we want them to be nearly equal, that is, we want their maximum value to be equal. The maximum betweenness for this dataset is near 400,000 and the maximum keyword similarity is 1. To compensate for this, we use a logarithmic scale for the betweenness centrality of authors. Please note that, as for now, we want the two types of similarities to be equal, but this may change in future due to evaluation of the algorithm. Also, future dyadic connections are considered, rather than multi-person cooperation. Doing so influences the way we compute the value of future cooperation. We will go into detail about this in the future work section.

In step four, the user receives a ranked list of peers in the network. We distinguish between two types of recommendations. That is, we can include or exclude existing co-authors in the recommendation. If the user chooses not to include existing co-authors, the user receives a list of only new candidate co-authors. We explicitly distinguish between these types of recommendations, as sometimes, people may prefer to write a new paper with existing co-authors rather than new co-authors, due to, for instance, trust, or time and location constraints. Figure 6.4 shows the user's welcome screen, which asks for the author's first and last name, and whether or not the authors wishes to include existing co-authors. Figure 6.5 shows an example of the resulting recommendation.

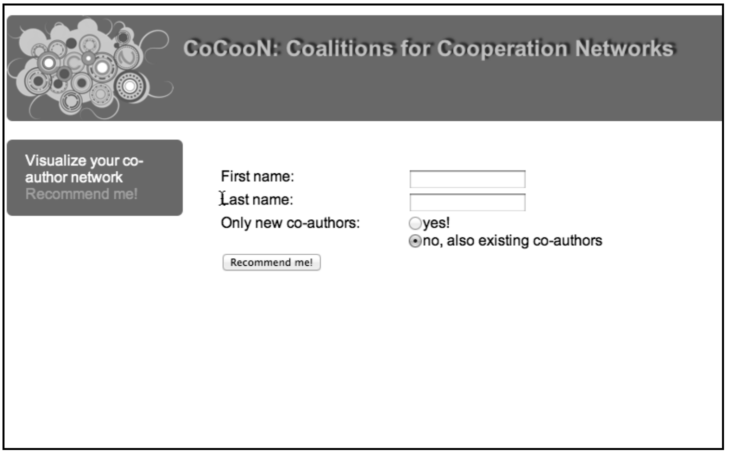


Figure 6.4. Example of the user's welcome screen.

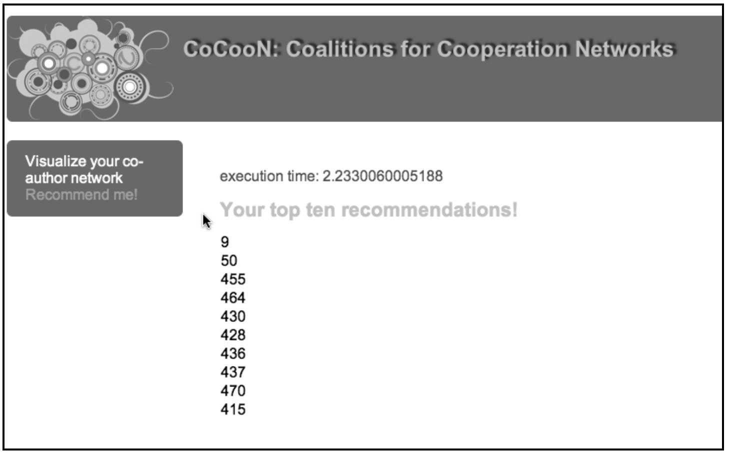


Figure 6.5. Example of the co-author recommendation. The candidate co-authors, denoted by numbers, are anonymised.

For clarification purposes, Table 6.3 provides a more formal representation of our algorithm, without going too much into detail about the computation of measures such as TFIDF and vector similarity.

Table 6.3. Recommendation algorithm.

Algorithm 2: Co-author recommendation based on betweenness centrality and keyword similarity

```
// create an empty stack for all peers in the network
W ← empty list;
// create empty stack of keywords
K[w] ← empty stack;
// create empty stack of TFIDF values per keyword and author
TFIDF[k,w] ← empty stack;
// create empty stack of vector similarity values for peers
VecSim[w], w ∈ W ← empty stack;
// create empty stack of utility values for peers
U[w], w ∈ W ← empty stack;
// extract all co-authors (see Table 2)
W ← extract coAuthors;
// create empty stack of peer's betweenness centrality
Cb[w] ← empty stack;

foreach peer w ∈ W do
   // save betweenness centrality
   push betweenness centrality of w → Cb[w];
   foreach submission s ∈ S do
      K[w] ← extract keywords;
      foreach keyword k ∈ K[w] do
         push compute TFIDF → TFIDF[k,w];
      end
   end
   push compute vector similarity to w → VecSim[w];
   push compute utility for w → U[w];
end
// sort the peers and their utility from high to low
sort U[w];
// repeat recommendation ten times
counter ← 0;
for counter < 10 do
   // recommend the peer
   recommendation = pop U[w];
   counter++;
end
```

6.2.3 Evaluation procedure

For the evaluation of the algorithm, we choose to conduct a pilot study. Since this is a first, and immature version of the recommendation engine, we aim to investigate the feasibility and identify possible improvements. We do not want to involve all potential participants from the sample (approximately 150 people), as they cannot be used for a later, large-scale evaluation due to prior experience with the system. Therefore, we contacted fifteen candidate participants to evaluate the two types of recommendation. The participants are all employees at the university that provided the DSpace dataset. They were invited by email, and were addressed personally. A total of ten participants responded positively.

Each of the fifteen participants received two sets of ten personal recommendations of future co-authors, sorted from high to low 'utility'. Set 1 was based on all authors that are present in the dataset. That is, we include the authors that the user has already written a paper with. This allows one to strengthen current ties in the network. However, some types of creativity are stimulated by connecting to new networks, or communities (Burt, 2004). Therefore, Set 2 solely consists of new future co-authors, people that the user has not yet written an article with.

For every co-author that was recommended, the participants had to assign a number ranging from 1 (bad) to 10 (good) to indicate the value of the recommendation. Further clarification said that our recommendation was based on 1) a person that has similar research interests, and 2) someone that has persuasive power, due to their occupation or network position. Thus, a 'good' recommendation should at least satisfy these two measures.

6.3 Results

Table 6.4 shows the results of the evaluation when current co-authors were included in the set of recommended future co-authors. The overall median is 7, which shows that the participants are in general quite positive towards the set of recommendations. As expected, the scores for the individual recommendations R1 to R10 gradually decrease, except for R8. Though, R8 shows an increase in score, but also high deviation.

Table 6.4. Results of the evaluation of recommendation Set 1, when current co-authors were included.

recommendation	***N***	***Mdn***	***SD***
Overall	10	7	2.68
R1	10	8.5	2.9
R2	10	8.5	1.5
R3	10	7	1.6
R4	10	7	1.7
R5	10	6.5	2.4
R6	10	5.5	3.2
R7	10	6.5	3.1
R8	10	8.5	3.3
R9	10	7	2.9
R10	10	6.5	2.9

Table 6.5 shows the results of the evaluation when current co-authors were excluded from the set of recommended future co-authors. The overall median is 6, which shows that the participants are in general quite neutral towards the set of recommendations. The scores for the individual recommendations R1 to R10 do not show a clear increase or decrease.

Table 6.5. Results of the evaluation of recommendation Set 2, when current co-authors were excluded.

recommendation	***N***	***Mdn***	***SD***
Overall	9	6	2.68
R1	9	6	2.00
R2	9	5.5	1.8
R3	9	5	2.3
R4	9	7	2.4
R5	9	6.5	2.5
R6	9	7.5	3
R7	9	6	2.7
R8	8	6	4
R9	9	4	2.8
R10	9	4.5	2.8

In response to the recommendation we sent, we received some statements from the participants:

1. *"Nothing really new, I also miss people I have obviously an overlap with like X, Y, Z, S, etc.."* This focuses on the functionality of the algorithm,

stating that its recall may be insufficient or that precision and recall may be unbalanced.

2. *"I don't know him."* This points to a lack of information provided by the system, or a lack of awareness of the user.
3. *"Some people I don't know, and others I do know, but I don't know what they do."* This points to a lack of information provided by the system, or a lack of awareness of the user.
4. *"He is now not active in research but has done work in the area I work in."* This points to lack of information within the system about active and inactive researchers.
5. *"He is now not very active in research."* This points to lack of information within the system about active and inactive researchers.
6. *"His research is now a bit different, games."* This points to user's preferences shifting in focus over time.

6.4 Discussion

In general, the results of this first test of our algorithm suggests that the participants are neutral to moderately positive about the recommendations that were generated. This leads us to believe that we are on the right track of combining network information with author similarity measures to recommend future co-authors.

The responses of the participants for Set 2 suggest that they are quite neutral toward the recommendations. Analysis of the responses shows that recommendations that are too distant from the target participant are regarded as pointless (statement 2 and 3). For example, one participant rated four out of ten recommendations with a 1, accompanied by the comment "I don't know him". This may point to lack of awareness, as observed in collaborative workspaces (Dourish & Bellotti, 1992; Reinhardt, Meier, Drachsler, & Sloep, 2011).

We may investigate how the participants rate recommendation of such 'distant persons' when they are presented how these people are linked to them, that is, the keywords that they have in common. In other words, explaining the workings of the recommender system may improve the user's perception (Herlocker, Konstan, Terveen, & Riedl, 2004; Sinha & Swearingen, 2002). Also, putting emphasis on the difference between the two sets of recommendations (Set 1 for strengthening bonds, Set 2 for creativity) may help in the adoption of recommendations.

The results for Set 1 indicate that participants are moderately positive about the recommendations of people that they already wrote a paper with. Though, some of the participants' comments indicate that the recommended people were not active in research anymore, or that the recommended person shifted focus over time (statement 4, 5 and 6). We could have gained higher ratings for this set of recommendations if we had compensated for changing preferences. Similar to “time-based discounting of ratings to account for drift in user interests” (Burke, 2002), we may perform time-based discounting of keyword-to-author relatedness.

6.4.1 Limitations

We need to take into account a number of limitations. First, we did not compensate for any misspelled author names or keywords. Sometimes, when people enter the names of their co-authors of their publication, they misspell the name, leading to two entries that point to the same person. To solve this, we would either have to compute the lexical similarity between a co-author’s name and the misspelled version of that co-author’s name, such as the Google similarity distance (Cilibrasi & Vitanyi, 2007) between them. Another option would be to manually search the database for any entries that are misspelled and save them in a thesaurus.

Second, people’s preferences can change over time. So can researchers’ interests. Throughout their scientific career, researchers often work in several universities or institutes, thereby inherently changing their focus, even if they keep working in the same research area. As a result of changing research interests, the keywords that researchers provided in publications from 2004 may be totally different than the keywords that they use in recent publications.

Third, and this follows partly from the previous point, time may influence our recommendation in another way. Researchers do not always stay in the same field of research, but may show up in recommendations based on their past publications. They may have even left research to work in business, or due to retirement. This severely influences the quality of our recommendations, as we will see in the results section. We will include this in future work.

6.5 Conclusion

In the present chapter we investigated how participants perceived utility-based recommendations of future co-authors. Expected utility originates from game theory and is especially useful to determine the expected value of a strategy, in this case a future co-authored paper. The main research question we asked ourselves was: *How well do participants perceive a recommendation that is based on keyword similarity and network information to be?* A small-scale evaluation was performed to determine the feasibility and receive intermediate feedback before we proceed with further development and a large-scale study. Neutral to moderately positive results indicate that the combination of network information (betweenness) and keyword similarity to recommend future co-authors is promising, but needs some improvements to maximize its potential.

The authors envisage two main points of improvement to the current recommender system. First, the current recommender system suggests dyadic connections, whereas co-authored papers often include more than two individuals.

The current algorithm is well suited to replace the dyad-based concept of utility by a solution concept that focuses on multi-person cooperation. We propose the use of coalition theory in general, and particularly the application of the Shapley value (Hart, 1987; Shapley, 1953) and the nucleolus (Kohlberg, 1971; Schmeidler, 1969) to value candidate cooperation partners, as noted by Sie *et al.* (2010b).

Secondly, we wish to account for drift in the users' research interests. Research interests change over time, and we need to compensate for this. Akin to Billsus and Pazzani (2000) and Pazzani (1999) that accounted for drift in user preferences, we need to give lower weight to keywords that were assigned to papers further back in time.

Thirdly, we wish to expand the dataset by including data from Mendeley (mendeley.com) and other DSpace publication databases, which are also freely accessible. This allows us to complete our network of candidate co-authors, and compute network information more precisely.

The next step in our research is to refine the system according to at least the above improvements. Furthermore, we aim to perform a large-scale evaluation of the recommender system.

CHAPTER 7

COCOON CORE: CO-author REcommendations based on Betweenness Centrality and Interest Similarity

This chapter presents the second version of the COCOON system, called CORE (CO-author REcommendation). Similar to the system in Chapter 6, it uses network position and interest similarity to recommend a future co-author to a target user. We made some significant improvements in the user interface of the system, and we added some extra features, such as an overview of researcher quality indices. The system was evaluated with a group of participants to investigate how they perceived the recommendations offered, and the system’s usability.

This chapter is based on: Sie, R.L.L., Van Engelen, B.J., Bitter-Rijpkema, M., & Sloep, P.B. (submitted). COCOON CORE: CO-author Recommendations based on Betweenness Centrality and Interest Similarity.

Abstract

When researchers are to write a new article, they often seek co-authors who are knowledgeable on the article's topic. However, they also strive for acceptance of their article. The current chapter presents the COCOON CORE tool that recommends candidate co-authors based on like-mindedness and power. Like-mindedness ensures that co-authors share a common ground, which is necessary for seamless cooperation. Powerful co-authors foster adoption of an article's research idea by the community. Two experiments were conducted, one focusing on the perceived quality of the recommendations that COCOON CORE generates and one focusing on the usability of COCOON CORE. Results indicate that participants perceive the recommendations moderately positively. Particularly, they value the recommendations that focus fully on finding influential peers and the recommendation in which they themselves can adjust the balance between finding influential peers and like-minded peers. Also, the usability of COCOON CORE is perceived to be moderately good.

7.1 Introduction

One of the main aims of a researcher, besides developing knowledge and understanding, is to strive for success and a solid reputation. Approaches to measure scientific successfulness such as the h-index (Hirsch, 2005) and the g-index (Egghe, 2006) exist, but it is still difficult for scholars (Linton, Tierney, and Walsh, 2011), journals (Gardner, Lowe, Moss, Maloney, & Cogliser, 2010), and agencies (Feuer, Towne, & Shavelson, 2002) to determine reputation and research success. Also, scholars are often unaware of the skills that they typically should attain to become successful. Indeed, being successful does not merely depend on performing high quality research, but also depends on the ability to reach out and convince others of the quality of a research idea. Researchers need to know what the main drivers for success are and they need to be made aware of these.

Lambiotte and Panzarasa (2009) draw attention to the fact that cohesive relationships in a topic-driven community foster researcher success. Articles need to be written, typically with co-authors, and these articles are subject to review. This requires a form of persuasion that involves knowledgeability and reputation. Leydesdorff and Wagner (2008) argue that power lies within a core group of network members. Also, they suggest that members in the periphery of the network can profit from more central members, consistent with Kotter's guiding coalition to lead organisational change (Kotter, 1996). Abbasi, Altmann and Hossain (2011) find that degree centrality, efficiency, tie strength and eigenvector centrality are indicators for a high g-index.

Current approaches to measure scientific success, such as the Hirsch spectrum tool (Franceschini & Maisano, 2010), take the distribution of the h-index of the journal's authors to measure the quality of a journal. Kim, Yoon and Crowcroft (2012) use network analysis to identify respected journals and proceedings. Particularly, they use node centrality and temporal analysis to provide insight into the emergence of scientific communities. SCImago (Falagas, Kouranos, Arencibia-Jorge, & Karageorgopoulos, 2008) provides an overview of a journal's impact, such as the h-index, number of citations, cited versus non-cited documents, etc.. The widely known Publish or Perish tool uses Google Scholar to measure an author's h-index or g-index (Harzing and Van der Wal, 2008). Yet, none of these tools aim at strategically bringing researchers into contact with co-authors to improve scientific success, as suggested by Lambiotte and Panzarasa (2010) and Leydesdorff and Wagner (2008).

The COCOON CORE tool aims to inform researchers about their personal quality and the strategically relevant researchers whom they should connect to. Its main functionality, presented in the current chapter, is the recommendation of candidate co-authors, which is based on two main principles: 1) co-author reputation (and power), which in turn is based on a central network position, and 2) interest similarity between a candidate co-author and the target user (common ground and shared intention), reflected by an overlap between keywords that two authors use to describe personal documents. It searches the open repository DSpace (http://www.dspace.org/) to aggregate and analyse the social network of individuals who co-authored documents. It has been built after the COCOON tool that generates co-author recommendations (Sie, Drachsler, Bitter-Rijpkema, & Sloep, in press). COCOON CORE caters to effective cooperation by finding candidate co-authors with a common ground and a shared intention. It does so by identifying peers in the network who have similar interests. Also, it caters to successful cooperation, by matching the target user with powerful, influential peers; peers who have authority, and are able to (indirectly) persuade others (e.g. reviewers).

The current chapter investigates what the opinion of the COCOON CORE user is toward the generated recommendations. As the recommendation calculation can be adjusted by the user by moving sliders, thus allowing one to focus on either influential peers or like-minded peers, it does not suffice to merely ask opinions about a recommendation that users can adjust themselves. To see how they value the two mechanisms, we also ask the users to focus fully on either mechanism. Hence, our research questions are as follows:

Research question 7.1: How do users value COCOON CORE's recommendation when they can adjust it to their personal preference?

Research question 7.2: How do users value COCOON CORE's recommendation when the algorithm fully focuses on influential peers?

Research question 7.3: How do users value COCOON CORE's recommendation when the algorithm fully focuses on like-minded peers?

Asking the user about the value of a recommendation can be influenced by the usability of the tool. To account for this, we conduct a standardised and widely established usability test called *SUS* (Brooke, 1996). The research question that follows from the usability test is as follows:

Research question 7.4: How do users experience the usability of COCOON CORE?

We start off the chapter with a discussion about the workflow of COCOON CORE, what data it uses and what calculations it performs (Section 7.2). We provide the method used to investigate the research questions (Section 7.3) and the results and discussion (Section 7.4). We draw this chapter to a close by providing our conclusion and a brief outlook on future improvements (Section 7.5).

7.2 COCOON CORE

7.2.1 Co-authorship network data

The data that we use to compute comes from a university's local publication database. The database, called DSpace (http://www.dspace.org), supports the open archives initiative, and its protocol, the OAI-PMH makes it possible for software to automatically extract metadata from the publications in the database. Documents are submitted to this database by (former) employees of the university. Table 1 provides an overview of the employees, departments, and publications that submitted to the database.

Table 7.1. Overview of the database (snapshot as of April 2012)

Publications	2924
Book chapters, articles and conference papers	1113
Presentations	904
Other	907
Authors	1,361
Keywords	3680
Departments	9

The data that we use to compute the centrality of co-authors is extracted from this database. For each document in the database, we extract its authors. These authors inherently form a co-authorship relationship. The aggregation of all authors of all publications forms a network of co-authors (Figure 7.1). As only (former) employees of the university submit documents to this database, the method of data collection is quite similar to that of an ego-centric network: a network as perceived form individuals' perspectives. Also, each document makes a clique; all authors of one document are interconnected through a bidirectional relationship.

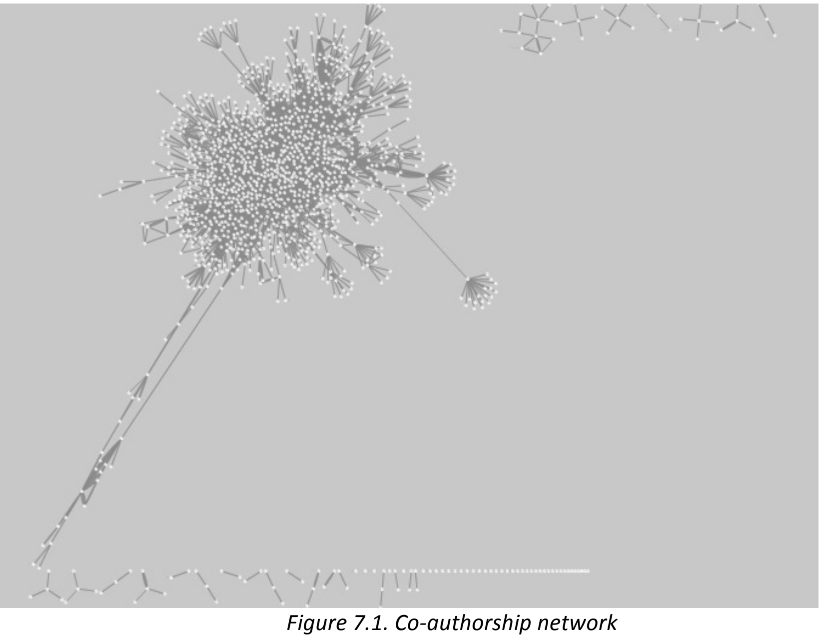


Figure 7.1. Co-authorship network

7.2.1 Calculations

The principal aim of COCOON CORE is to recommend candidate co-authors. Its algorithm employs two types of calculations to arrive at the recommendation. First, for every author in the social network, it computes the power, or reputation of an author; to what extent other authors are dependent on the target author in terms of disseminating ideas within the network. It does so by taking the number of times a target author is on the shortest path between any two other authors in the network relative to the total number of shortest paths, also known as *betweenness centrality* (Freeman, 1977; Brandes, 1994).

Second, the algorithm computes similarity between authors. High similarity, in gender for instance, is found to be an indicator for good relationships (Ibarra, 1992), and this is supported by research on homophily and friendships (Lazarsfeld & Merton, 1954; McPherson, Smith-Lovin, & Cook, 2001). Stahl (2005) argues that cooperation between any two authors be guided by a common ground. To measure similarity, we first have to identify individuals within the network. For each author, we look at her submissions and the keywords that she has used in these submissions, and construct a keyword vector. The distance between authors' keyword vectors defines the similarity between authors (*vector similarity).*

7.2.3 Recommendation workflow

The workflow of COCOON CORE is depicted in Figure 7.2. The workflow commences with user Polly, who wants to write a new paper. A new paper requires a topic, so Polly starts defining the paper's topic or main research idea.

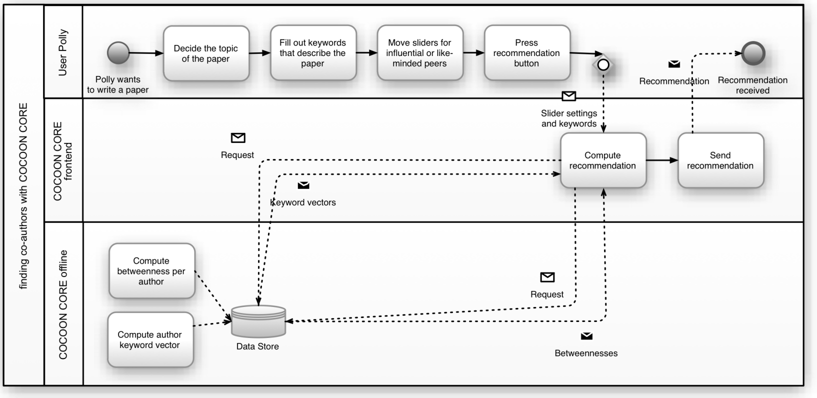


Figure 7.2. Workflow for a COCOON CORE recommendation.

Next, Polly fills out the keywords that describe her paper's topic (Figure 7.3) and decides whether COCOON CORE should favour like-minded peers or influential peers. For instance, if Polly is exploring a topic in which she has relatively low authority, she may decide to focus on finding influential, powerful peers. She does so by moving the sliders to her preference. Figure 7.3 shows slider settings that favour like-minded peers (bottom slider), which reflects the situation that Polly already has some authority in the research field. Finally, she presses the button 'GIVE RECOMMENDATION' and COCOON CORE starts computing a recommendation. Thus, the main user interactions with COCOON CORE comprise 1) filling out keywords, 2) moving sliders to preference, and 3) pressing the 'give recommendation' button.

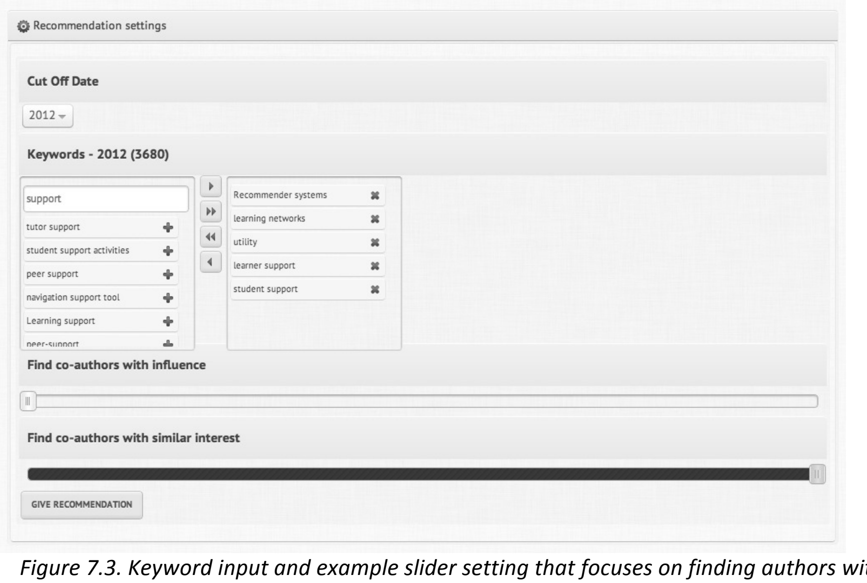


Figure 7.3. Keyword input and example slider setting that focuses on finding authors with similar interest.

As indicated, Polly put in keywords that describe the topic of the new paper. These keywords, together with keywords that already exist in her personal keyword vector, are used to compute and find authors that are like-minded. Also, the slider settings define how much focus should be put on the similarity between authors by the recommendation engine. In detail, this is achieved by sending a request to the COCOON CORE backend, which already computed the keyword vector. The backend replies by sending the author keyword vectors, and now the similarity between authors can be computed.

Next, a request for influential peers is sent to the backend data store. The backend data store replies by sending back the betweenness centrality of each author. The slider setting now define to what extent the betweenness (influential peers) and keyword similarity (like-minded peers) should be taken into account to compute the final score per peer. For instance, if the slider for influential peers is set to 20, then the normalised betweenness score (between 0 and 1) will be multiplied by 0.20, whereas the normalised keyword similarity will be multiplied by 0.80. A typical recommendation result is shown in Figure 7.4. The authors (Figure 7.4, column 2) are sorted by their calculated score (Figure 7.4, column 1). Besides, authors can be sorted using their betweenness (Figure 7.4, columns 3 and 4) and keyword similarity (Figure 7.4, column 5).

Recommended Autors - 2012

Show 10 entries

Search:

Score	Author	Betweenness 2012	Normalized Betweenness 2012	keyword similarity 2012
89.8622494477	Peter Dolog (dspace)	24429	0.817780111755	0.952517416291
89.3679355411	Paul A. Kirschner (dspace)	93773	0.926653085784	0.871696868496
89.2675333072	Wim Westera (dspace)	15245	0.779615301282	0.968048687599
88.9495411085	George Sielis (dspace)	11098	0.753917908376	0.979880412892
87.6052139781	Dirk Börner (dspace)	6852	0.714886637642	0.983495807873
87.5434463186	Renate De Groot (dspace)	53329	0.880970778275	0.87174358646
86.1654937193	Bert Hoogveld (dspace)	34806	0.846434420705	0.871801948186
86.1614753084	Davinia Hernández-Leo (dspace)	34813	0.84645069731	0.8717241236
85.7448849309	Wolfgang Greller (dspace)	17874	0.792492507114	0.900753077438
85.4038929965	Argyris Kouloumbis (dspace)	3712	0.665272822431	0.979883001654

Figure 7.4. COCOON CORE recommendation result. The first column shows the final score, the second shows the recommended authors and their DSpace link. The third, fourth and fifth column show intermediate computation results.

7.3 Methodology

7.3.1 Participants

Participants in this experiment were 23 employees from the investigated university that hosts the DSpace repository in question (*N*=23, *total population*=89). All participants were selected based on their use of DSpace; they were active as a researcher and had uploaded at least one document. The group consisted of 13 male and 10 female participants with a tenure ranging from 1 to 35 years (*M* = 9.48; *SD* = 7.84). Their occupation ranged from PhD researcher to full professor. Participation was voluntary and beside homemade pastry, no inducement was offered.

7.3.2 Materials

7.3.2.1 'Find your co-author' task

The participants had to perform three tasks for which they had to evaluate the recommendation corresponding to the research question in point (cf. Section 1). First, they were asked to set the slider for influence to 100 per cent. The slider for interest similarity was automatically set to zero per cent. Second, they were asked to set the slider for interest similarity to 100 per cent. The slider for influence was automatically set to zero per cent. Finally, they were asked to adjust both sliders to their individual liking.

7.3.2.2 Task Instruction

Before the start of the task, participants were provided with a detailed briefing document that showed the basic functionality of the tool. The briefing showed how to login, how the dashboard functioned, and how they should put in keywords in order to generate a recommendation. One of the researchers was present either in person or online to support remote participants, but no serious issues arose. The task instruction lasted 10 minutes in total.

7.3.2.3 Recommendation questionnaire
Participants were asked to answer three questions on a five-point Likert scale (1 = very bad, 5 = very good), corresponding to the three tasks for their individual recommendation and for the default user recommendation, respectively (Appendix B).

7.3.2.4 System Usability Scale (SUS)
Next to testing the quality of the recommendations generated by COCOON CORE, we wanted to receive feedback on its user-friendliness (research question 7.4). The standardised and widely used System Usability Scale (SUS) was used to evaluate the usability of COCOON CORE. SUS conforms to the ergonomics of human-computer interaction DIN EN ISO 9241, part 11. Overall, it measures the perceived usability of the tool at hand and sub-scales include usability (questions 1-3 and 5-9, Appendix C) and learnability (questions 4 and 10). SUS is an industry standard with over 5000 users and 500 reported studies. In detail, it contains ten questions that can be answered using a five-point Likert scale (1=strongly disagree, 5=strongly agree)(Appendix B). The final SUS score ranges from 0 (bad usability) to 100 (good usability) points. On average, systems evaluated using the SUS usability test score 68 points.

7.3.3 Design and procedure
Each participant has a different profile in the DSpace repository, which is dependent on the frequency of uploads and the keywords that they use to describe the document. For reasons of comparability, the experiment therefore included an evaluation of a recommendation for a default user's profile in DSpace besides the evaluation for the participants' individual profile. The default user profile consisted of one the author's profiles, whose articles were present in the database as well.

A between-subjects design was used, in which participants had to perform the three tasks for a default user (*D*), and for themselves (*S*). The main reason for this was to overcome a sequence bias in evaluation of COCOON CORE. Group 1 started with task *D*, and subsequently performed task *S*. Group 2 started with task *S*, and subsequently performed task *D* (Table 2). The participants were randomly assigned to Group 1: *DS* (*N*=12) or Group 2: *SD* (*N*=11).

Table 7.2. Task sequence for two participant groups

Group 1: *DS* condition (*N* = 12)	Default user recommendation *D*	Individual recommendation *S*
Group 2: *SD* condition (*N* = 12)	Individual recommendation *S*	Default user recommendation *D*

7.3.4 Data analyses

Difference between groups were tested for statistical significance using an independent samples t-test for each of the questions regarding the individual and default user recommendation (six in total). No significant difference between these groups would mean that there is no effect in the sequence in which these tasks are performed.

Note that the rating is reversed for each subsequent question in the SUS questionnaire; the odd-numbered questions' scores are calculated by the scale position minus one (e.g. 5 is a good rating, and results in a score of 4), and the even-numbered questions' scores are calculated by 5 minus the scale position the participant gave (e.g. 1 is a good score, and results in a score of 4). Next, the scores are multiplied by 25 to arrive at a scale between zero and 100.

7.4 Results and discussion

7.4.1 Recommendation questionnaire

Table 7.3 shows the significance tests for the answers to each of the six questions regarding the recommendations (Figures 7.5 and 7.6). It shows that the two groups do not significantly differ from one another for each and every question. This means that there is no sequence effect between the two groups. In other words, it did not matter which recommendation task was given first, the individual recommendation task or the default user recommendation task. For example, Levene's test shows that with respect to question 1b, the two groups do not significantly differ ($t(22) = .924$, $p < 0.05$).

Table 7.3. Results of Levene's independent samples t-test.

question	t	df	Sig.
1a	.000	22	.737
1b	-,924	22	.371
1c	-1.999	22	.653
2a	3.924	22	.177
2b	-.705	22	.707
2c	.240	22	.736

N=24

The medians for each recommendation question (Figure 7.5) show that participants are moderately positive toward the recommendations generated.
With respect to the individual recommendations, we can conclude that participants score the recommendation in which the influence slider is set to 100 (research question 7.1, recommendation 1a) scores moderately positively. The individual recommendation in which the interest similarity slider is set to 100 (research question 7.2, recommendation 1b) scores neutral. The individual recommendation in which participants can adjust the sliders themselves (research question 7.3, recommendation 1c) scores moderately positively. This implies that participants

particularly value the recommendations that either fully focus on finding influential peers, or the recommendation that they can adjust to their personal preference.

When compared with the default user's recommendations (recommendations 2a, 2b, and 2c), the ratings of the individual recommendations score slightly higher. For example, the individual recommendation in which the influence slider is set to 100 (question 1a) scores equally high compared to the same recommendation for the default user (question 2a). Also, the individual recommendation in which similarity is set to 100 (question 1b) scores equally high compared to the same recommendation for the default user (question 2b). However, individual recommendation in which the sliders are set to personal preference (question 1c) scores slightly higher than the same recommendation for the default user (question 2c). This discrepancy may be due to the users' lack of familiarity with the default user's work. For example, we quote one participant: "*harder to judge, as this is not really my topic, than when searching with my keywords. But looks good.*"

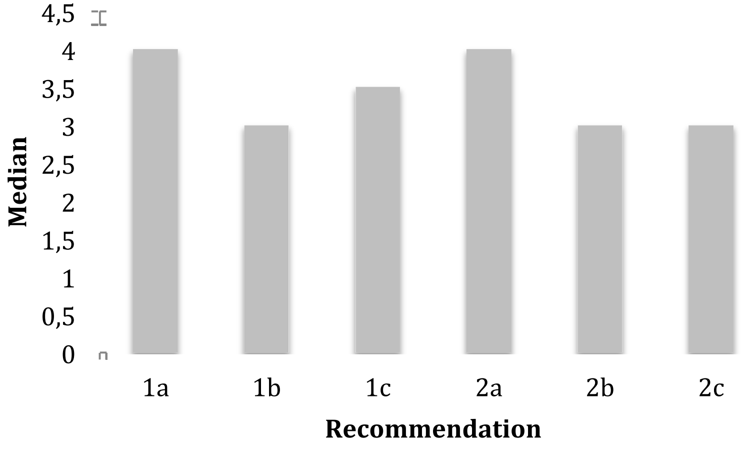


Figure 7.5. Median for each recommendation question.

A closer look at the proportion of responses (Figure 7.6) reveals that participants are especially positive toward the recommendation that focuses entirely on influential peers (1a and 2a) and the recommendation in which participants could set the sliders to their personal preference (1c).

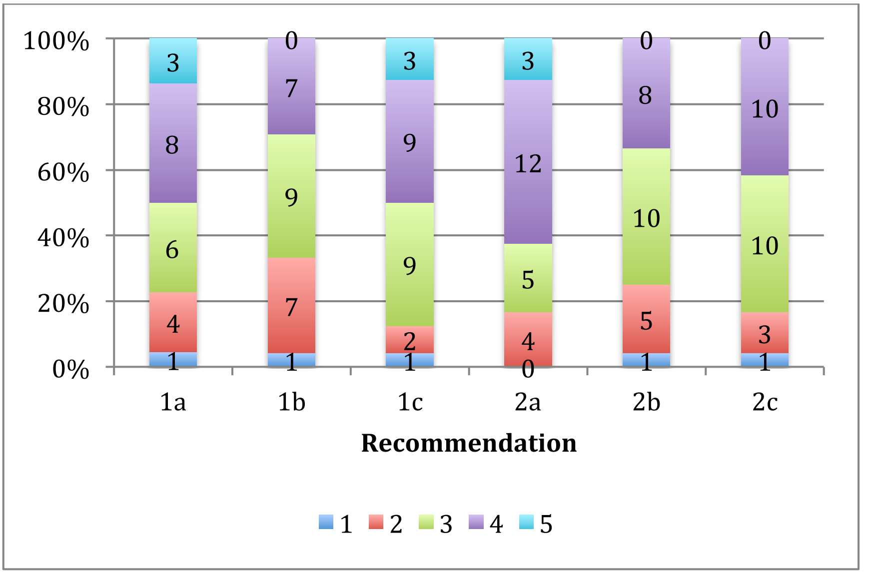


Figure 7.6. Proportion of responses for each recommendation question.

Thus, a recommendation that is based on successfulness and effective cooperation satisfies the users to a moderately positive extent. Regarded from a more algorithmic level, a combination of betweenness centrality to identify powerful, influential peers in the network, and vector similarity to identify like-minded peers satisfies the participants, and shows to have potential.

Our recommendation results are partly in contrast with research by Abbasi, Altmann and Hossain (2011), who found no significant effect of betweenness centrality on the g-index. This disparity can be explained as follows. COCOON CORE focuses on successful and effective cooperation, rather than increasing the g-index. In other words, COCOON CORE aims at increasing acceptance for papers, but also agreeable cooperation between co-authors. Numerous papers are rejected, and the reason for this is not always clear. Naturally, a paper should be rejected on the basis of lack of quality, and this could have been due to a lack of common ground among authors. The g-index is based on accepted papers that are highly cited, and does not reflect the actual successfulness of cooperation between authors. Furthermore, the nature of Abbasi *et al.*'s g-index is different from the current study, which measures user satisfaction and usability.

7.4.2 System Usability Scale (SUS)

The SUS usability test brings forward that COCOON CORE scores fairly positively on a normalized scale of 0 to 100 (*Mdn* = 67.50, Table 7.4). At a confidence interval of 95% and a sample size of 24, this means that the average usability value is likely to fluctuate between 57.57 and 72.42.

Table 7.4. Summary of System Usability Scale (SUS).

Measure	Value
Min	25
M	65.27
GM	65.25
Mdn	67.50
Max	90
95% confidence interval	57.57 - 72.42

N=24

Figure 7.7 shows that participants are especially positive about the learnability of COCOON CORE (questions 4 and 10, Figures 7.7 and 7.8), for instance not needing a technical person to use COCOON CORE (question 4). Also, when looking at the proportions of responses (Figure 7.8), participants think that there are few inconsistencies in COCOON CORE (question 6) and that COCOON CORE is not unnecessarily complex (question 2).

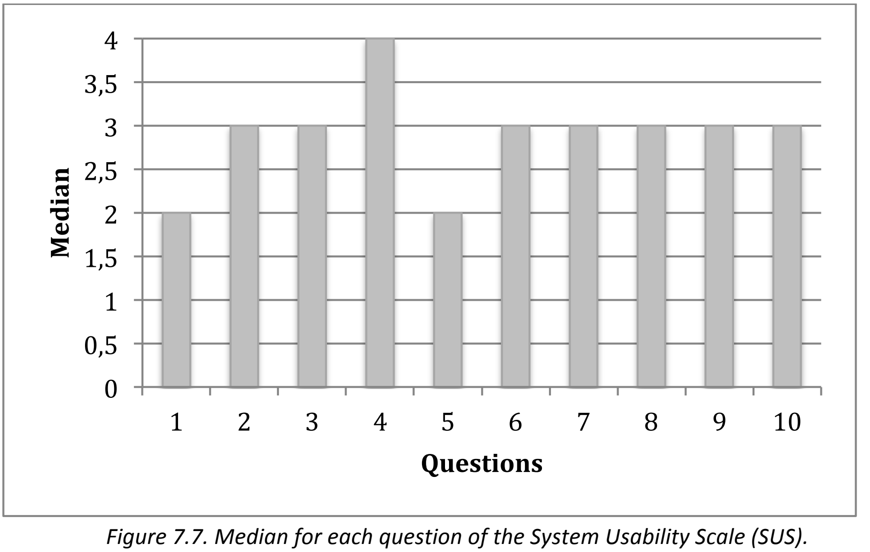


Figure 7.7. Median for each question of the System Usability Scale (SUS).

A closer look at Figure 7.8 reveals that the most notable shortcoming lies in the integration of several functions (question 5). The proportion of responses for question 5 show that fourteen out of 24 participants (58%) rated the integration of functions neutral to negative. This was expected, as functions such as author metrics and recommendations were distributed among several pages. Nevertheless, a future version of COCOON CORE should focus more on the integration, or at least the visual integration of functionality.

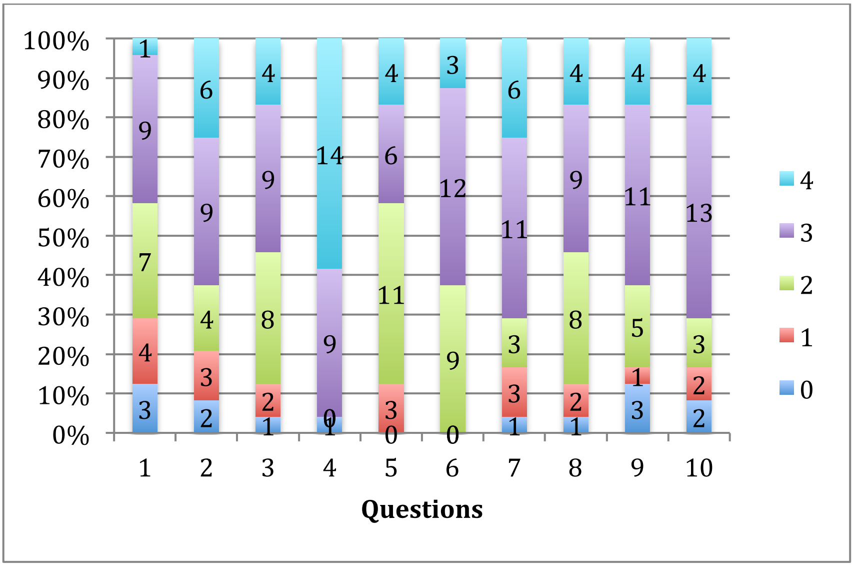


Figure 7.8. Proportion of responses for each question of the System Usability Scale (SUS).

7.5 Conclusion

The tool presented here (COCOON CORE) recommends co-authors based on power and influence of peer co-authors (betweenness centrality), and a common ground between prospective co-authors (keyword vector similarity). It strives to increase the chance of paper acceptance, and pleasant cooperation among co-authors, respectively. The nature of research questions was two-fold. Firstly, we measured the perceived quality of recommendations, both from participants' individual perspective and default user's perspective. Secondly, we measured the usability of COCOON CORE by means of the standardised and widely used System Usability Scale (SUS), arguing that a low usability would influence the quality score negatively.

Participants perceive the usability of COCOON CORE as moderately positive. Especially the learnability of COCOON CORE (no technical assistance required) scores high and users do not face too much inconsistency. Therefore, no negative influence on the appreciation of co-author recommendations is expected. That said, next to an overall improvement of the usability, improvements should be made with respect to the integration of functionality, such as the author metrics and the recommendation engine.

Crucially, a combination of betweenness centrality and keyword vector similarity, respectively, is found to be useful. This result points to the usefulness of COCOON CORE as a co-author recommender. Note that this is partly out of line with earlier research in which no significant effect was found for betweenness centrality and the g-index. However, this study aimed at perceived quality of a recommendation system (user satisfaction), rather than measuring researcher quality based on longitudinal data, thus explaining the discrepancy.

Future work should focus on longitudinal analysis of the successfulness of these recommendations. That is, it should investigate whether recommended co-authorships lead to higher researcher performance. To make such analyses possible, the authors plan to implement additional functionality that allows COCOON CORE users to directly or indirectly (through gatekeepers or the system as a mediator) approach a candidate co-author.

Acknowledgements

The authors thank Dr. Lora Aroyo from the VU University Amsterdam for her insightful comments during the design and implementation phases of COCOON CORE.

CHAPTER 8

General Discussion

This chapter draws the results reported in the previous chapters together and attempts to paint an integral picture of what has been achieved as well as what questions, urgent or not so urgent, are still outstanding.

8.1 Introduction

In today's global economy it has become key that we cooperate. In 2012, the World Economic Forum (WEF) released a report about the need for collaboration to drive economic growth (Antoniou, Arkless, Bedford, Bochniarz et al., 2012). They exemplified a number of good practices. Firstly, they mention cooperation through the pooling of talent, or talent mobility. Two of the main issues that currently hold back talent mobility are gaps in information and gaps in skills. Both can be resolved by bringing talents into contact with the right peers in their network; peers that have the complementary knowledge that is required for a talent to increase mobility. Also, the information gaps are due to a lack of awareness; employers are not aware of what individuals have on offer, and individuals are not aware of what possibilities lie ahead of them.

Secondly, the WEF calls for *effective* collaboration, that is, "Building the Right 'Muscles'." To clarify, they argue that collaboration must be guided by 1) a common ground, 2) shared intention, 3) strong governance, 4) hard evidence of results, and 5) continuous assessment of progress and results. In this thesis, we addressed three of these requirements. We argue for the need for a common ground (homophily) between individuals to guide cooperation. Moreover, we call for a shared intention as part of a successful 'coalition' in cooperation networks. Finally, we contend that reputation, status, and authority (strong governance) may guide the successful implementation of innovative (research) ideas. **It is not collaboration by itself that is important, it is whom you collaborate with[3].**

In this chapter, we will look back on the progress we have made in our attempt to enhance cooperation in networks. We will do so by revisiting the main research questions that we posed in Chapter 1.4: 1) what factors influence cooperation between networked individuals and 2) how can we persuade individuals to cooperate so that their idea will be accepted or implemented. Next, we will inventory our results and explore what the practical implications of our results are. Finally, we present our research vision for the upcoming years. This includes work in finding the right peers in cooperation networks, but also other, less apparent directions, such as enhancing creativity itself, and empowering network members by improving their cognition of the network.

[3] Please note that to stay within the terminology of the WEF, we mention collaboration. However, we argue that the majority of collaboration is in fact cooperation, because often partners may have shared intentions but also have distinct goals. See Chapter 1.2 for further explanation.

8.2 Key contributions

8.2.1 Theory

The initial focus of this thesis was on the finding out which factors influenced cooperation networks, to inform the design and implementation of our simulations and support tool. We conducted two distinct experiments to answer the question which factors influence cooperation networks (*research question 1*). In experiment one, we asked professional learners – professional learners - how they perceive their personal learning using their social network (Chapter 2, research question 1a). In experiment two, we asked two groups of experts to discuss what factors in their expert opinion influence cooperation networks (Chapter 3, research question 1b). In both experiments, we sorted the initial set of factors to arrive at core clusters of factors that influence cooperation networks.

In experiment 1 (Chapter 2), we found that the viewpoints of learners toward their personal professional networked learning (*research question 1a*) can be divided into seven core clusters: sharing, motivation, perceived value of the network, feedback, personal learning, trust and support, and peer value and characteristics. Perceived value of the network along with peer value and characteristics are the reason *why* learners engage in networked learning. Sharing and trust and support are key to *how* learners should learn via their networks. *What* learners learn mainly results in personal learning, and is driven by feedback given by peers.

Also, the way professional learners engage in networked learning has changed slightly now that we are using online social tools. Intuitively, one would think that social bookmarking tools such as Delicious.com or other ways of capturing knowledge (Wikis, podcasts, blogs, scoop.it) would be the main means of networked learning, but they were rarely mentioned. Rather, networked learners use email, face-to-face contact; their only 'concession' to the modern Internet is their usage of Twitter to connect to peers.

As experiment 1 primarily focused on the network practitioners themselves, there was a need for a higher level, less subjective perspective of domain experts. Experiment 2 (Chapter 3) focused on the question what factors influence cooperation in networks (*research question 1b*) according to experts. We asked two groups of experts - one heterogeneous, one homogeneous – to generate and discuss such factors. Based on these expert discussions, we found that there are four core clusters of factors that influence cooperation in networks: personality and motivation, diversity, effective cooperation, and management and interpersonal relationships.

We elicited knowledge from three participant groups from distinct domains. Having three distinct participant groups allowed us to come up with more general findings. Firstly, we asked professional learners to provide their take on learning via their network (Chapter 2). Secondly, a heterogeneous group of experts from such diverse domains as psychology, innovation and game theory gave their view on what factors influence cooperation in networks. Thirdly, a heterogeneous group of experts on learning networks offered their perspective on factors influencing cooperation in networks.

According to both the domain experts and the learners motivation is an important aspect of cooperation in networks. Learners and experts agreed on the cluster motivation to be a core influencing factor. Also, trust was mentioned by the individual groups as a crucial factor. Moreover, the experts rated trust among the most important factors that influence cooperation in networks. This is consistent with research by Rusman et al. (2009) on trust in virtual teams.

Learners in a network are primarily goaded into self-interested action: receiving feedback, support and the value that the network and its members have on offer (Chapter 2). In Chapter 3, the experts agree on trustworthy relationships, shared goals and joint interests. In other words, cooperation networks thrive on reciprocity. The results of Chapter 2 address a unidirectional learning connection, rather than a bidirectional, reciprocal relationship. Therefore, the results of Chapter 2 can only partially be extended to cooperation networks and learning networks in general. This may be due to the nature of the questions that we asked the participants of the experiment in Chapter 2, which were mainly focused on learning, rather than teaching through the network.

We contend that peer value and characteristics can be identified for cooperation networks, and trust and sharing can be catered to by tailored software that finds a peer to cooperate with. However, this should not be just any peer. How to find this peer was the main reason for research question 2, which focused on persuading individuals.

8.2.2 Simulation

In Chapter 4 and 5, we investigated how the factors that we identified in Chapters 2 and 3 - augmented with factors that a literature search revealed - relate to one another (*research question 1c*). We implemented two models that simulate how these factors influence cooperative behaviour of individuals in an innovation network. The first simulation (Chapter 4) showed that agents with low power can loaf and rely on agents with high power to have their idea implemented. Conversely, agents with high power can use agents with low power to reach the necessary majority to have their idea implemented, also known as social loafing (Latané et al., 1979; Karau & Williams, 1993; Liden et al., 2004; Chidambaram & Tung, 2005).

The simulation in Chapter 5 showed that the average betweenness centrality[4] of a winning coalition is highly predictive of the average power of a winning coalition: as betweenness increases, the average power of a winning coalition decreases. At first sight, this may seem odd. However, 'average power of a winning coalition' implies that a coalition has already won. Thus, if you have high betweenness, then it is easier to stand out (lower average power) and have success in implementing your idea. This is consistent with theories about the strength of weak ties. Weak ties can lead to a higher betweenness, and high betweenness is associated with being influential (Brass, 1984). Having weak ties can make you more creative (Burt, 2004), as the weakly tied peers offer you a variety of viewpoints different from your own. Thus, high betweenness and high average betweenness in a coalition can help you implement your innovative idea.

The multi-agent simulations in Chapters 4 and 5 were based mainly on literature study and the two experiments in Chapters 2 and 3. A common approach to simulation in many AI studies is to use only literature data to build a simulation model, and such a model aims to simulate real-life behaviour by means of a simplified version of reality. A simplified model does not capture each and every factor that influences behaviour in real life. Ideally, we would want a complex, multi-level model of each and every factor that influences behaviour, see how the results feed back to the model, which then influences behaviour in a different way, and so on. Common practice shows that often when we make a model more complex, it loses its predictive capabilities. This is the main reason why we tried to triangulate the factors from literature with knowledge from experts and practitioners. The best way we could describe these simulations is that they have an *explorative* character, albeit based on triangulated data.

To draw more accurate conclusions, we need to base our simulations on existing data, rather than on literature. The simulations in Chapter 4 and 5 were carried out before we had laid our hands on the research publication dataset that we presented in Chapter 6 and 7. Future research should focus on designing and testing a model that is based on these real world data. Such models can also be used to predict future evolution of the behaviour exhibited in this dataset. We must note, however, that models that resemble the real world in detailed often do not simulate the real world faithfully. Finally, to arrive at general conclusions, we should design and simulate models using distinct datasets to compare whether the factors in this simulation model hold.

---

[4] High betweenness centrality means that a network member – the co-author – is often on the shortest path between any two other network members. Being on the shortest path between two other members means that the co-author can influence the knowledge that passes through him.

8.2.3 Researcher support

In Chapter 6, we presented a researcher support tool (COCOON) that aims to assist researchers in their search for a new co-author when they plan to write a new article. Many approaches to monitor researcher success exist, such as the h-index (Hirsch, 2005) and the g-index (Egghe, 2006), but none so far have focused on supporting the researcher in finding strategic partnerships, even though this has been suggested some time ago already (Leydesdorff & Wagner, 2008; Lambiotte & Panzarasa, 2010). Hence, COCOON aimed at assisting the researcher in finding the right future co-author, rather than just any co-author. Every researcher has ideas, and most of them believe their research idea is worth publishing. However, in practice, not all good ideas are always implemented. Therefore, any researcher, in fact any innovating individual, needs support to implement his or her ideas. COCOON does so by recommending them key individuals.

We believe that a future co-author who has the ability to persuade others should meet two main requirements (research question 2a). Firstly, the co-author should be an authority in the field. That is, the co-author should have a form of power, in this case the power over information flow, that is, the power to influence what knowledge is spread, and to whom. In social network analysis terms, such a powerful co-author is associated with a high betweenness centrality. Our simulations in Chapter 5 emphasised the importance of betweenness centrality for the acceptance of an idea.

Secondly, the future co-author should be *knowledgeable*. Being knowledgeable on a topic adds up to one's success rate when trying to persuade others, next to knowing about the target that is to be persuaded, and knowledge about persuasion itself (Friedstad & Wright, 1994). Moreover, Wesch (2009) makes an important distinction in how we should handle the current digital revolution in which knowledge is growing for ever to the point of overloading people; we should become able to handle knowledge (*knowledge-able*) instead of just having knowledge (knowledgeable). That is, we should focus on where to find knowledge, and how to filter out the right knowledge. Indeed, this is what a network member with high betweenness centrality should be able to do. Also, individuals that have something in common (homophily) are more likely to cooperate well (Ibarra, 1992). Thus, while striving for more persuasive power through knowledgeability and knowledge-ability, at the same time COCOON increases the probability of successful cooperation between two individuals.

COCOON followed a two-pronged recommendation approach. Firstly, in an institutional setting we retrieved co-authored submissions from an open archive-based database called DSpace. It yielded a network of individual researchers who cooperated on creating media such journal articles, presentations, conference papers and project deliverables. Secondly, we computed the similarity between candidate co-authors from the keywords they supply when uploading a submission.

The weighted average of these two metrics led to a ranked list of recommended co-authors.

In more detail, we created two ranked lists of recommended co-authors to see if there is a difference in perception between recommended co-authors that users know, and recommended co-authors that users do not know yet (*research question 2b*). Thus, each user received two lists of ten recommended authors: one that included both existing and new, possibly unknown co-authors, and one that included only existing co-authors.

The results show that users favour the list of existing co-authors over the one that also includes new co-authors. Indeed, some of the user comments and ratings revealed that they were unfamiliar with certain recommended co-authors, resulting in low evaluation scores for the 'unknown' recommendations of co-authors.

In Chapter 7, we presented a new version of the same co-author recommendation tool: COCOON CORE. Its main improvements lie in its dashboard functionality. After user login, COCOON CORE shows a dashboard in which personalized recommendations can be obtained. The user herself can put in keywords for the paper to be written, and emphasise either finding co-authors with similar interest or finding influential co-authors.

We conducted an evaluation session with a group of researchers from the university that hosts the DSpace database. We specifically focused on this group of researchers, because they 1) were active researchers, and 2) they submitted their work in the database. During evaluation, we addressed four research subquestions that together address this thesis' research question 2c, based on the configurations of the tool. First, we looked if the researchers agreed on the tool's choice of a set of people with a similar interests. Second, we investigated to what extent researchers agree on its choice of a set of people that have influential power. Third, we studied how researchers perceived recommendations that were generated based on their own preferences wit respect to influential peers and similar peers. Finally, we studied how the researchers perceived the tool's usability. This meant they could set their own preferences for the search options and choose their own keywords. This is in line with the practice of a researcher who wants to write a new article with co-authors. First, the topic is defined (e.g. keywords), and then the researcher starts looking for knowledgeable and perhaps powerful or authoritative peers.

The results show that the COCOON CORE users rate the co-author recommendations moderately positively, particularly when they modify the sliders for finding influential peers and like-minded peers themselves. Thus, a combination of betweenness centrality and keyword vector similarity is found to be useful when recommending future co-authors. Besides, COCOON CORE users also perceive its usability moderately positively. Specifically the learnability of the tool scores high and users do not face too much inconsistency in terms of functionality. Consequently, we contend that we are well on our way to developing a tool that can bring together researchers such that they can cooperate well and be successful at the same time.

The way researchers co-author publications does not always reflect their actual contribution to the paper, thereby posing a challenge for the definition of the strength of a tie between two individuals. For instance, in PhD research, it is common for the PhD candidate to include the names of the daily supervisor and the overall supervisor in an article, because they had their say while conducting the experiment and during the writing of the article. We cannot, however, provide definitive percentage estimates on the extent of their contributions; for instance, that the first author contributed 70 per cent to a co-written paper, the second author contributed twenty per cent, and the third author contributed ten per cent. In fact, the contribution of the individual authors may vary per paper, but may also per author as they have distinct personalities. We argue for a method that can bypass an individual's contribution to a paper in defining the quality of a researcher. It seems that current approaches in co-citation analysis (Fisichella, Herder, Marenzi, & Nejdl, 2010) and output metrics such as the h-index (Hirsch, 2005) and g-index (Egghe, 2006) can already form quite an elaborate picture of a researcher's quality.

8.2.4 Methodology

Along the way, we made three, minor but in our view useful contributions to research methodology. Firstly, we designed an online environment to conduct the *eDelphi*, an electronic version of the Delphi methodology. The eDelphi environment helps a researcher to elicit knowledge from a group of experts – often dispersed – to let the group reach consensus, and to analyse the results. It provides several 'dashboard views' that illustrate the productivity of the group of experts as a whole and the productivity per individual in the group.

Secondly, we created a new type of methodology to elicit information or opinions from subjects. It is based on the brainstorming technique in that it comprises an idea generation phase without discussion, because often ideas are lost during offline creative sessions with co-workers due to production blocking (Nijstad, Stroebe, & Lodewijkx, 2003). Besides, it is conducted via Twitter, which allows for quick and dispersed participation. Moreover, online tweets - the ideas - can be easily aggregated by using a 'hashtag' and automatic backup software such as twapperkeeper (http://twapperkeeper.com/). The advantages of Twitter as a medium and brainstorm as a technique resulted in the methodology name *Tweetstorm*, which is a merger of the two.

Finally, we created an environment that collects learning network data from an ego-perspective, the *COCOON PLN identification tool*. It consists of a form that asks participants for the peers that they learnt from, and how they connected to the peers. The learning relationships constitute a personal learning network, and when we gather enough of these relationships, we can analyse it to identify key ‘tutors’ in the network, or key ‘tools’ that are used to learn from peers. Also, the data may be used to recommend valuable peers in the network that one can learn from. The aggregation and combination of network data from several contexts can, ultimately, be used to compare the characteristics of the social networks to arrive at general conclusions about and interventions in these networks.

8.4 Some practical implications and suggestions for future research

A day without an idea is a day wasted. Without ideas, we could not have conducted the research in this thesis. All research starts out with an idea and, preferably, has some practical consequences. Therefore, this section will not only point to a number of possible practical consequences but also provides some thoughts on how the research in this thesis should continue.

8.4.1 Practical implications

The COCOON PLN identification tool is a valuable instrument to discover from whom learners learn and to analyse the learning networks that the tool yields. Before we proceed with retrieving additional data about learning networks, a few enhancements should be made to release its full potential. We first need to refine the form’s questions. The results in Chapter 2 show that some answers may be sorted into categories like ‘microblogging’, videoconferencing and bookmarking. Furthermore, by asking *what learners learn*, we can make sense of the topics that may or may not drive communities or clusters of learners. Moreover and similar to the co-author recommendations, we can use the network information to *recommend valuable peers* in the learning network. Finally, this tool is not only limited to eliciting learning relationships. With some minor adjustments, we can allow for moderator-generated questions and answers to yield other, domain-specific networks. That is, we can, for instance, ask participants whom they innovate with, to see how an innovation network looks like from an individual’s perspective. We can also ask participants whom they trust, to yield a trust network. In other words, we can open up the COCOON PLN identification tool to uses other than merely in service of learning networks.

COCOON CORE is a tool that has the potential to be incorporated in institutional repositories such as DSpace and then give recommendations. In principle, it could also be linked to, for example, Mendeley and give recommendations that go beyond institutional boundaries (assuming this is in actual fact feasible). For COCOON CORE, to provide even better recommendations, a number of improvements come to mind, that we also intend to implement in future releases. Firstly, we plan to improve the user experience by more apparent integration of services as called for in Chapter 7. Next, we plan on adding new features, such as author profiles, adding social media handles, adding the g-index to the author performance dashboard, and integrating the co-author recommendation with the graph visualisation. The advantage of adding author profiles can, for instance, increase the chance of adoption, as the user comments in Chapter 6 indicate. Also, author profiles can be shared among candidate teams to increase trust, as suggested by Berlanga, Rusman, Bitter-Rijpkema and Sloep (2009).

Secondly, and following from the previous enhancement, we plan to improve the way we store and access the data. That is, we plan to enhance information retrieval and analysis of the co-author and keyword graphs by using more elaborate indexing and caching mechanisms. Such indexing optimization can, for instance, be performed by creating 'MySQL views' to cache queries into data files on the server that COCOON CORE runs on. We may decide to store data using *semantic web techniques* such as RDF or OWL. A semantic database allows for automated reasoning, sense making and enrichment of data by using open, linked data. Naturally, in that case, optimisations for MySQL queries will not work. The most recent version of SPARQL (Prud'hommeaux & Seaborne, 2012), an RDF store in some ways similar to a MySQL database, allows for querying of a path depth. This means that mining one's social network can easily and efficiently be limited to, say, three hops in the network.

Thirdly, the knowledge of a researcher is not only represented by her publications. Nowadays, researchers use blogs, podcasts, wikis, and social media such as Twitter, Facebook and LinkedIn to reach out and share their thoughts. A next version of COCOON CORE should use social media information to optimize the 'profiling' of the researcher. In other words, we can use Twitter and Facebook posts to determine the latest interest of the researcher more precisely (cf. Drachsler, 2009). Also, we can use social media to determine which of the keywords that the researcher uses are currently trending, by using so-called *sentiment analysis* (Pang & Lee, 2008). Naturally, we can apply this to analyse trending topics in other types of networks as well, such as innovation networks, or learning networks.

8.4.2 Future research

Two of the main problems that we tried to solve in this thesis are the lack of awareness and the availability of only bounded rationality. We assume that these problems affect decision making, but do they actually affect researchers' decisions? As far as we know, there are no methods available to determine a lack of awareness or the presence of a bounded rationality. Therefore, we argue for methods to measure these indicators of *network cognition*. One way to discern lack of awareness, for instance, may be comparison of an ego-network with a complete network. That is, we compare the individual's perspective on one's own network with a network that is based on facts, such as email traffic, to see if the individual can pinpoint all her contacts. Sie, Ullmann, Rajagopal and Cela (submitted) mention the use of Near Field Communication to monitor contact moments between individuals, in order to capture a complete network of relationships.

Based on extensive literature review in the learning domain, Sie, Ullmann, Rajagopal and Cela (submitted) conclude that intervention and simulation are two major gaps in the domain of *social network analysis for learning*. Research should continue on using social network analysis to help learners, innovators and researchers identify key peers in their network that can help them advance. This comprises using social network analysis to intervene in the daily lives of learners, innovators and researchers, but also to inform multi-agent simulations of social networks to predict future behaviour. The DSpace data about co-authorships (Chapters 6 and 7) could well act as a starting point to building a simulation model, by performing multiple regression analysis on this data. A simulation model that is rooted in real world data may provide a more accurate perspective on how behaviour resulting from a social network analysis-driven system will evolve.

In the domain of *recommender systems*, time-drift is a common problem in determining user profiles. Users' preferences change over time, so recommending a book based on books bought between 2008 and 2012 may raise some eyebrows, whereas a recommendation based on books bought in 2012 may yield a higher chance of approval. We argue that time-dependence also plays a role in determining social networking behaviour, and thus calls for applications that perform an intervention based on the current dynamics of the network, rather than 'old' dynamics of the network.

Although work has been done in the representation of social network by means of semantic web formats such as RDF, the field has not taken off in this direction yet. We argue for the use of *semantic web* representations of cooperation networks such as learning networks, innovation networks and research networks, to make sense of the data and perform automated reasoning on the data. Peter Mika (2009, p. 163-182) has commenced similar work by visualising and analysing research communities of interest.

Building on the work of Mika, Ereteo *et al.* (2009) have created SPARQL queries to measure centrality for specific types of relationships, such as 'friend', 'family' or 'colleague'. In this way, we can more accurately analyse the relationships and positions of individuals in social subnetworks. For example, we can analyse what the degree centrality of an individual within a family is, by only calculating degree centrality over the 'family'-relationships. We argue for a similar approach in the storage and analysis of cooperation networks such as learning networks. The centrality of an individual may differ from topic to topic. *When we distinguish between knowledge topics and types of relationship, we can more accurately bring together peers that can learn from one another.* Analogous to spiders that are primarily subsocial, but cooperate when put together with genetic kin (Ruch, Heinrich, Bilde, & Schneider, 2009), working or learning together may be boosted by bringing together like-minded or otherwise related individuals.

When researchers work on an article, they rarely write an article on their own. It is common for three or more authors to work together on a single article, but current recommendation approaches do not address the need for 'group recommendations'. Solution concepts from *game theory* such as the Shapley value and the nucleolus formalize the value of groups. Moreover, these solution concepts can be used in such a way that they account for common maxims 'two heads are better than one' or 'the whole is more than the sum of its individual parts' that state that cooperating groups can outperform nominal groups. In other words, current recommendation algorithms recommend cooperation between dyads, whereas cooperation often takes place between more than two individuals. Recommendation algorithms that allow for group valuation are needed, and the Shapley value and the nucleolus can do this. The simulations in this thesis made a first attempt to using game theoretic solution concepts.

8.3 In conclusion

In the introductory chapter to this thesis, we laid out four types of problems that individuals encounter when they engage in cooperation through their social network. Analysing all four types of problems and suggesting ways to overcome them proved to be too much for one thesis. In the end, we focused mainly on solving the interpersonal and intrapersonal problems and paid little attention to procedural, structural and exogenous problems.

With respect to the interpersonal perspective, we tried to solve problems such as the lack of awareness of whom one can co-author an article with by making people aware of the valuable peers in their network. Besides, this approach aimed at decreasing information overload by offering only a limited number of recommended co-authors. We also tried to compensate for the bounded rationality that individuals experience, their inability to solve the kind of complex judgement that is needed to efficiently and effectively value the peers in their network.

With respect to intrapersonal perspective, we tried to develop a tool that fosters reciprocity and aims to use self-interest in a productive way by showing the value of cooperation to both parties involved in cooperation. Naturally, we tried to recommend valuable peers to individuals, but the recommendation algorithm was based on similarity of interests as well. The former makes for solving the intrapersonal problems, the latter aims to foster reciprocity. Through a simulation we showed that agents with high power can profit from low-power agents, because the low power agents can account for the necessary majority that one needs to persuade other members in the network.

By addressing the interpersonal and intrapersonal problems that may arise in cooperation networks, we hope to have brought cooperation in networks closer by.

Innovation networks, research networks and learning networks stand to profit from this.

References

References

Abbasi, A., Altmann, J., & Hossain, L. (2011). Identifying the effects of co-authorship networks on the performance of scholars: A correlation and regression analysis of performance measures and social network analysis measures. *Journal of Informetrics*, *5*(4), 594-607. Elsevier Ltd. doi:10.1016/j.joi.2011.05.007

Abdallah, S., & Lesser, V. (2004). Organization-based cooperative coalition formation. *Proceedings. IEEE/WIC/ACM International Conference on Intelligent Agent Technology, 2004. (IAT 2004).*, 162-168. Ieee. doi:10.1109/IAT.2004.1342939

Acquisti, A., Sierhuis, M., Clancey, W., & Bradshaw, J. M. (2002). Agent Based Modeling of Collaboration and Work Practices Onboard the International Space Station. *Proceedings of the 11th Conference on Computer-Generated Forces and Behavior Representation.*

Adamides, E. D., & Karacapilidis, N. (2006). Information technology support for the knowledge and social processes of innovation management. *Technovation*, *26*(1), 50-59. doi:10.1016/j.technovation.2004.07.019

Adidas. (2009). Adidas "JABULANI" Official Match Ball of the 2010 FIFA World Cup^{TM}. Retrieved from http://www.soccerballworld.com/jabulani_2010.htm

Aiken, M., Krosp, J., Shiran, A., & Martin, J. (1994). Electronic brainstorming in small and large groups. *Information & Management*, *27*(3), 141-149. doi:10.1016/0378-7206(94)90042-6

Albino, V., Carbonara, N., & Giannoccaro, I. (2006). Innovation in industrial districts: An agent-based simulation model. *International Journal of Production Economics*, *104*(1), 30-45. doi:10.1016/j.ijpe.2004.12.023

Allen, M. P., & Tildesley, D. J. (1999). *Computer Simulation of Liquids* (p. 385). Oxford University Press, USA.

Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology*, *80*(3), 260-267. doi:10.1037//0022-0663.80.3.260

Antoniou, C., Arkless, D., Bedford, P., Bochniarz, H., Brown, J. F., Burns, M. M., Clarke, T., et al. (2012). *Talent Mobility Good Practices Collaboration at the Core of Driving Economic Growth* (p. 1-110).

Arrow, H., McGrath, J. E., & Berdahl, J. L. (2000). *Small groups as complex systems: formation, coordination, development and adaptation* (p. 336). Thousand Oaks, California, USA: SAGE Publications, Inc.

Attwell, G. (2007). Personal Learning Environments - the future of eLearning? *Lifelong Learning*, *2*(January), 1-8.

Aumann, R. J. (1959). Acceptable points in general cooperative n-person games. *The Theory of Games IV, Annals of Mathematics Studies*, *40*, 287-324.

Auranen, O., & Nieminen, M. (2010). University research funding and publication performance—An international comparison. *Research Policy, 39*(6), 822-834. Elsevier B.V. doi:10.1016/j.respol.2010.03.003

Aviv, R., & Ravid, G. (2005). Reciprocity Analysis of Online Learning Networks. *Journal of Asynchronous Learning Networks*, *9*(4).

Axelrod, R. (1984). *The Evolution of Cooperation*. Basic Books, Inc.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, *211*(4489), 1390. AAAS. Retrieved from http://www.sciencemag.org/cgi/content/abstract/sci;211/4489/1390

Bacharach, S. B. (2005). DIVERSITY AND HOMOPHILY AT WORK : SUPPORTIVE RELATIONS AMONG WHITE AND AFRICAN-AMERICAN PEERS. *Dana*, *48*(4), 619-644.

Bandura, A. (1977). *Social Learning Theory*. New York, NY, USA: General Learning Press.

Barnard, C. I. (1968). *The Functions of the executive* (p. 334).

Beham, G., Kump, B., Ley, T., & Lindstaedt, S. (2010). Recommending knowledgeable people in a work-integrated learning system. *Procedia Computer Science*, *1*(2), 2783-2792. doi:10.1016/j.procs.2010.08.003

Belbin, M., & Belbin, R. (1996). *Team roles at work* (2nd Editio., p. 162). Butterworth-Heinemann Oxford. Retrieved from http://www.lavoisier.fr/notice/frMTOALX3HXOQALO.html

Berendt, B., & Kralisch, A. (2007). From World-Wide-Web Mining to Worldwide Webmining : Understanding People’s Diversity for Effective Knowledge Discovery. *LNAI 4737*, 102-121.

Berlanga, A. J., Bitter-Rijpkema, M. E., Brouns, F., & Sloep, P. B. (2008a). On the Importance of Personal Profiles to Enhance Social Interaction in Learning Networks. *Proceedings of Web Based Communities Conference (WEBC 2008)* (pp. 55-62). Amsterdam, The Netherlands: IADIS Press.

Berlanga, A. J., Sloep, P. B., Kester, L., Brouns, F., Rosmalen, P. V., & Koper, R. (2008b). Ad hoc transient communities: towards fostering knowledge sharing in learning networks. *International Journal of Learning Technology*, *3*(4), 443. doi:10.1504/IJLT.2008.019378

Berlanga, A., Rusman, E., Bitter-Rijpkema, M., & Sloep, P. (2009). Guidelines to Foster Interaction in Online Communities. *LEARNING NETWORK SERVICES FOR PROFESSIONAL DEVELOPMENT* (pp. 27-42). Springer Berlin / Heidelberg. doi:10.1007/978-3-642-00978-5_3

Bhattacharyya, S., & Ohlsson, S. (2010). Social creativity as a function of agent cognition and network properties: A computer model. *Social Networks*, *32*(4), 263-278. doi:10.1016/j.socnet.2010.04.001

Billett, S. (2004). Workplace participatory practices: Conceptualising workplaces as learning environments. *Journal of Workplace Learning*, *16*(6), 312-324. doi:10.1108/13665620410550295

Billsus, D., & Pazzani, M. J. (2000). User Modeling for Adaptive News Access. *User Modeling and User-Adapted Interaction*, 147-180.

Bitter-Rijpkema, M., Martens, R., & Jochems, W. (2002). Supporting knowledge elicitation for learning in virtual teams. *Educational Technology & Society*, *5*(2).

Boden, M. A. (2004). *The creative mind: Myths and mechanisms*. London: Routledge. Routledge.

Boland Jr, R. J., & Tenkasi, R. V. (1995). Perspective making and perspective taking in communities of knowing. *Organization science*, *6*(4), 350–372. JSTOR. Retrieved from http://www.jstor.org/stable/2634993

References

Borgatti, S. P., & Cross, R. (2003). A Relational View of Information Seeking and Learning in Social Networks. *Management Science*, *49*(4), 432-445

Brandes, U. (1994). A Faster Algorithm for Betweenness Centrality. *Journal of Mathematical Sociology*, *25*, 163-177.

Brass, D. J. (1984). Being in the Right Place: Structural Analysis of Individual Influence in an Organizaton. *Adminstrative Science Quarterly*, *29*(4), 518-539.

Brass, D. J., Galaskiewicz, J., Greve, H. R., & Tsai, W. (2004). Taking Stock of Networks and Organizations: A Multilevel Perspective. *The Academy of Management Journal*, *47*(6), 795. doi:10.2307/20159624

Boud, D., & Middleton, H. (2003). Learning from others at work: communities of practice and informal learning. *Journal of Workplace Learning*, 15(5), 194-202. doi:10.1108/13665620310483895

Brooke, J. (1996). SUS - A quick and dirty usability scale. *Usability evaluation in industry*, *189*, 1-8.

Brueckner, S. A., & Van Dyke Parunak, H. (2003). Resource-Aware Exploration of the Emergent Dynamics of Simulated Systems. *AAMAS'03* (pp. 718-788). Melbourne, Australia: ACM.

Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction*, 1-29. Retrieved from http://www.springerlink.com/index/N881136032U8K111.pdf

Burkhardt, M. E., & Brass, D. J. (1990). Changing Patterns or Patterns of Change: The effects of a change in technology on social network structure and power. *Administrative Science Quarterly*, *35*(1), 104-127.

Burnstein, E., & Vinokur, A. (1977). Persuasive argumentation and social comparison as determinants of attitude polarization. *Journal of Experimental Social Psychology*, 13(4), 315-332. doi:10.1016/0022-1031(77)90002-6

Burt, R. S. (2004). Structural Holes and Good Ideas. *American Journal of Sociology*, *110*(2), 349-399. doi:10.1086/421787

Cassiman, B., & Veugelers, R. (2006). In search of complementarity in innovation strategy: Internal R&D and external knowledge acquisition. *Management Science*, *52*(1), 68.

Chesbrough, H. W. (2003). The Era of Open Innovation. *MIT Sloan Management Review*, 35-42.

Cheung, C. M. K., & Lee, M. K. O. (2010). A theoretical model of intentional social action in online social networks. *Decision Support Systems*, *49*(1), 24-30. doi:10.1016/j.dss.2009.12.006

Chidambaram, L., & Tung, L. L. (2005). Is Out of Sight, Out of Mind? An Empirical Study of Social Loafing in Technology-supported systems. *Information Systems, 16*(2), 149-168.

Cilibrasi, R., & Vitanyi, P. (2007). The Google Similarity Distance. *IEEE Transactions on Knowledge and Data Engineering*, *19*(3), 370-383. doi:10.1109/TKDE.2007.48

Clancey, W.J., Sierhuis, M., Kaskiris, C., & Van Hoof, R. (2003). Advantages of Brahms for specifying and implementing a multiagent human-robotic exploration system. *Proc. FLAIRS-2003*, 2002(Figure 1), 7–11.

Colman, A. M., Pulford, B. D., & Rose, J. (2008). Collective rationality in interactive decisions: evidence for team reasoning. *Acta psychologica, 128*(2), 387-97. doi:10.1016/j.actpsy.2007.08.003

Colman, A. M., & Pulford, B. D. (2012). Problems and Pseudo-Problems in Understanding Cooperation in Social Dilemmas. *Psychological Inquiry, 23*(1), 39-47. doi:10.1080/1047840X.2012.658003

Crano, W. D. (1995). Attitude strength and vested interest. In J. A. Petty & R. E. Krosnick (Eds.), *Attitude strength: Antecedents and Consequences* (pp. 131-158). Lawrence Erlbaum Associates, Inc.

Crano, W. D., & Prislin, R. (1995). Components of Vested Interest and Attitude-Behavior Consistency. *Basic and Applied Social Psychology, 17*(1-2), 1-21. doi:10.1080/01973533.1995.9646129

Cross, R., Parker, A., Prusak, L., & Borgatti, S. P. (2001). Knowing What We Know: Supporting Knowledge Creation and Sharing in Social Networks. *Organizational Dynamics, 30*(2), 100-120.

Cyert, R. M., & March, J. G. (2005). *A behavioral theory of the firm*. Blackwell.

Dall'Asta, L., Marsili, M., & Pin, P. (2012). Collaboration in social networks. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(12), 4395-400. doi:10.1073/pnas.1105757109

Das, T. K., & Teng, B.-S. (1997). Sustaining Strategic Alliances: Options and Guidelines. *Journal of General* Manage*ment*, *22*(4), 49-64.

Davidsson, P. (2002). Agent Based Social Simulation : A Computer Science View. *Journal of Artificial Societies and Social Simulation*, *5*(1), 4-8.

Davis, K., & Alexander, K. L. (2009). A Focused Research Agenda for Family and Consumer Sciences Education : Results of a Policy Delphi Study. *Journal of Family & Consumer Sciences Education*, *27*(2), 30-44.

Dawson, S., Bakharia, A., & Heathcote, E. (2010). SNAPP : Realising the affordances of real-time SNA within networked learning environments. *Proceedings of the 7th International Conference on Networked Learning* (pp. 125-133).

De Choudhury, M., Sundaram, H., John, A., & Seligmann, D. (2008). Dynamic prediction of communication flow using social context. *Proceedings of the nineteenth ACM conference on Hypertext and hypermedia - HT '08*, 49. New York, New York, USA: ACM Press. doi:10.1145/1379092.1379105

Defays, D. (1977). An efficient algorithm for a complete link method. *The Computer Journal*, *20*(4), 364-366.

Dekker, J. M., & Kingma, J. (1999). 'Managen' van leernetwerken. In J. M. Dekker (Ed.), *Netwerkend Leren (Opleiders in Organisaties: Capita Selecta vol 38)*. Deventer: Kluwer.

Denning, P. J. (2012). The Profession of IT, the Idea Idea: What if practices rather than ideas are the main source of innovation? *Communications of the ACM*, *55*(3), 30-32.

Dignum, V. (2002). A Knowledge Sharing Model for Peer Collaboration in the Non-Life Insurance Domain. *Proceedings of German Workshop on Experience Management*.

References

Dillenbourg, P. (1999). What do you mean by “collaborative learning”? In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and Computational Approaches* (Vol. 1, pp. 1-15).

Distler, O., Behrens, F., Pittrow, D., Huscher, D., Denton, C. P., Foeldvari, I., Humbert, M., et al. (2008). Defining appropriate outcome measures in pulmonary arterial hypertension related to systemic sclerosis: a Delphi consensus study with cluster analysis. *Arthritis and rheumatism*, *59*(6), 867-75. doi:10.1002/art.23718

Dourish, P., & Bellotti, V. (1992). Awareness and coordination in shared workspaces. *Proceedings of the 1992 ACM conference on Computer-supported cooperative work - CSCW ’92*, 107-114. New York, New York, USA: ACM Press. doi:10.1145/143457.143468

Downes, S. (2010). Learning networks and connective knowledge. In H. H. Yang & S. C.-Y. Yuen (Eds.), *Collective Intelligence and E-Learning 2.0: Implications of Web-Based Communities and Networking* (Vol. 28, pp. 1-26). IRMA International. doi:10.4018/978-1-60566-729-4.ch001

Drachsler, H. (2009). *Navigation Support for Learners in Informal Learning Networks*. SIKS Dissertation Series. Retrieved from http://dspace.learningnetworks.org/handle/1820/2084

Dunbar, R. I. M. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences, 16*, 681-735.

Ebner, M., Schön, S., Taraghi, B., & Drachsler, H. (2011). First steps towards an integration of a Personal Learning Environment at university level. In R. Kwan et al. (Ed.), *ENHANCING LEARNING THROUGH TECHNOLOGY. EDUCATION UNPLUGGED: MOBILE TECHNOLOGIES AND WEB 2.0* (Communicat., Vol. 177, pp. 22-36). Berlin, Heidelberg: Springer Verlag. doi:10.1007/978-3-642-22383-9_3

Egghe, L. (2006). Theory and practise of the g-index. *Scientometrics*, *1*(January 2006), 1-31.

Ensminger, D. C., & Surry, D. W. (2008). Relative ranking of conditions that facilitate innovation implementation in the USA. *Australasian Journal of Educational Technology*, *24*(5), 611-626.

Eraut, M. (2004). Informal learning in the workplace. *Studies in Continuing Education*, *26*(2), 247-273. doi:10.1080/158037042000225245

Ereteo, G., Gandon, F., Corby, O., & Buffa, M. (2009). Semantic Social Network Analysis. *Web Science WebSci’09* (pp. 1-5). Athens, Greece. Retrieved from http://hal.inria.fr/inria-00378174/

Falagas, M. E., Kouranos, V. D., Arencibia-Jorge, R., & Karageorgopoulos, D. E. (2008). Comparison of SCImago journal rank indicator with journal impact factor. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, *22*(8), 2623-8. doi:10.1096/fj.08-107938

Fernández, A., & Gómez, S. (2008). Solving Non-Uniqueness in Agglomerative Hierarchical Clustering Using Multidendrograms. *Journal of Classification*, 25, 43-65. doi:10.1007/s00357-008-

Fetter, S., Berlanga, A., & Sloep, P. (2009). Enhancing the Social Capital of Learning Communities by Using an Ad Hoc Transient Communities Service. *Advances in Web Based Learning–ICWL 2009*, (1), 150–157. Springer. Retrieved from http://www.springerlink.com/index/46u521w616q14355.pdf

Fetter, S., Berlanga, A. J., & Sloep, P. (2010). Fostering Social Capital in a Learning Network: Laying the Groundwork for a Peer-Support Service. *International Journal of Learning Technology*, 5(4), 388-400.

Feuer, M. J., Towne, L., & Shavelson, R. J. (2002). Scientific Culture and Educational Research. *Educational Researcher*, *31*(8), 4-14. doi:10.3102/0013189X031008004

Finke, R. A. (1996). Imagery, creativity, and emergent structure. *Consciousness and Cognition*, *5*(3), 381-393. doi:10.1006/ccog.1996.0024

Finkelstein, M. J., & Lacelle-Peterson, M. W. (1992). New and junior faculty: A review of the literature. *New Directions for Teaching and Learning* (50), 5-14. doi:10.1002/tl.37219925003

Fisichella, M., Herder, E., Marenzi, I., & Nejdl, W. (2010). Who are you working with ? - Visualizing TEL Research Communities - Collecting Co-Citation Data. *World Conference on Educational Media & Technology (EdMedia)*.

Fogel, J., & Nehmad, E. (2009). Internet social network communities: Risk taking, trust, and privacy concerns. *Computers in Human Behavior*, *25*(1), 153-160. Elsevier Ltd. doi:10.1016/j.chb.2008.08.006

Follett, M. P., & Metcalf, H. C. (2003). *Dynamic administration: The collected papers of Mary Parker Follett: Early sociology of management and organizations* (Vol. 3). Psychology Press.

Franceschini, F., & Maisano, D. (2010). The Hirsch spectrum: A novel tool for analyzing scientific journals. *Journal of Informetrics*, *4*(1), 64-73. doi:10.1016/j.joi.2009.08.003

Freeman, L. C. (1977). A Set of Measures of Centrality Based on Betweenness. *Sociometry*, *40*(1), 35-41.

Friedstad, M., & Wright, P. (1994). The Persuasion Knowledge Model: How People Cope With Persuasion Attempts. *Journal of Consumer Research*, *21*(June), 1-31.

Gallupe, R. B., Dennis, A. R., Cooper, W. H., Valacich, J. S., Bastianutti, L. M., & Nunamaker, J. F. (1992). Electronic Brainstorming and Group Size. *Academy of Management Journal*, *35*(2), 350-369. doi:10.2307/256377

Gardner, W. L., Lowe, K. B., Moss, T. W., Mahoney, K. T., & Cogliser, C. C. (2010). Scholarly leadership of the study of leadership: A review of The Leadership Quarterly's second decade, 2000–2009. *The Leadership Quarterly*, *21*(6), 922-958. Elsevier Inc. doi:10.1016/j.leaqua.2010.10.003

Gerstick, C. J. G., Bartunek, J. M., Dutton, J. E. (2000). Learning from academia: The importance of relationships in professional life. *The Academy of Management Journal*, 43(6), 1026-1044. doi:10.2307/1556333

Gigerenzer, G., & Selten, R. (2001). *Bounded rationality: The adaptive toolbox*. MIT press.

References

Gilbert, N., Pyka, A., & Ahrweiler, P. (2001). Innovation Networks - A Simulation Approach. *Journal of Artificial Societies and Social Simulation*, *4*(3). Retrieved from http://jasss.soc.surrey.ac.uk/4/3/8.html

Granovetter, M. S. (1973). The strength of weak ties. *American journal of sociology*, *78*(6), 1360–1380. JSTOR. Retrieved from http://www.jstor.org/stable/2776392

Gulbrandsen, M., & Smeby, J.-C. (2005). Industry funding and university professors' research performance. *Research Policy, 34*(6), 932-950. doi:10.1016/j.respol.2005.05.004

Hanak, E., & Rueben, K. (2006). Funding Innovations for California Infrastructure : Promises and Pitfalls. *Public Policy*, (March).

Harary, F., Norman, R. Z., & Cartwright, D. (1965). *Structural models: An introduction to the theory of directed graphs*. New York, NY, USA: Wiley.

Hart, S. (1987). Shapley value. *The New Palgrave*, 318-320.

Harzing, A., & van der Wal, R. (2008). Google Scholar as a new source for citation analysis. *Ethics in Science and Environmental Politics*, *8*, 61-73. doi:10.3354/esep00076

Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of advanced nursing*, *32*(4), 1008-15. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11095242

Hauser, C., Tappeiner, G., & Walde, J. (2007). The learning region: the impact of social capital and weak ties on innovation. *Regional Studies*, *41*, 75-88.

Haythornthwaite, C., & De Laat, M. (2010). Social Networks and Learning Networks : Using social network perspectives to understand social learning. In L. Dirckinck-Holmfeld, V. Hodgson, C. Jones, M. de Laat, D. McConnell, & T. Ryberg (Eds.), *Proceedings of the 7th International Conference on Networked Learning* (pp. 183-190).

Herlocker, J. L., Konstan, J. a., Terveen, L. G., & Riedl, J. T. (2004). Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems*, 22(1), 5-53. doi:10.1145/963770.963772

Hill, R. A., & Dunbar, R. I. M. (2003). Social network size in humans. *Human Nature, 14*(1), 53-72.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(46), 16569-72. doi:10.1073/pnas.0507655102

Hollingshead, A. B., Fulk, J., & Monge, P. (2002) Fostering intranet knowledge sharing: An integration of transactive memory and public goods approaches. In S. Keisler & P. Hines (Eds) *Distributed work: New Research on Working Accross Distance Using Technology* (pp. 335-355). Cambridge, MA: MIT Press.

Ibarra, H. (1992). Homophily and Differential Returns : Sex Differences in Network Structure and Access in an Advertising Firm. *Science*, *37*(3), 422-447.

Ibarra, H. (1993a). Network Centrality, Power, and Innovation Involvement: Determinants of Technical and Administrative Roles. *Academy of Management Journal*, 36(3), 471-501.

Ibarra, H. (1993b). Personal networks of women and minorities in management: A conceptual framework. *Academy of Management Review*, 18(1), 56-87. doi:10.2307/258823

Ibarra, H., Kilduff, M., & Tsai, W. (2005). Zooming In and Out : Connecting Individuals and Collectivities at the Frontier of Organizational Network Research. *Organization Science*, *16*(4), 359-371.

idSpace. (2012). idSpace homepage. Retrieved May 1, 2012, from http://www.idspace-project.org/

Isenberg, D. J. (1986). Group Polarization : A Critical Review and Meta-Analysis, 50(6).

Janis, I. L. (1982). *Groupthink: Psychological studies of policy decisions and fiascoes*. Houghton Mifflin Boston, MA.

Jarvenpaa, S. L., & Leidner, D. E. (1998). Communication and Trust in Global Virtual Teams. *JCMC*, *3*(4).

Jensen, M., & Roy, A. (2008). Staging Exchange Partner Choices: When Do Status and Reputation Matter? *The Academy of Management Journal*, *51*(3), 495. doi:10.2307/20159523

Jones, C. R., Ferreday, D., & Hodgson, V. (2008). Networked learning a relational approach: weak and strong ties. *Journal of Computer Assisted Learning*, *24*(2), 90-102. doi:10.1111/j.1365-2729.2007.00271.x

Jonker, C. M., Robu, V., & Treur, J. (2007). An agent architecture for multi-attribute negotiation using incomplete preference information. *Autonomous Agents and Multi-Agent Systems, 15*(2), 221-252. doi:10.1007/s10458-006-9009-y

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263-292.

Karau, S. J., & Williams, K. D. (1993). Social Loafing: A Meta-Analytic Review and Theoretical Integration. *JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY*, *65*, 681-681.

Kau, J. B., & Rubin, P. H. (1979). Self-interest, ideology, and logrolling in congressional voting. *Journal of Law and Economics*, 22(2), 365-384.

Keltner, D., Kleef, G. A. V., Chen, S., & Kraus, M. W. (2008). A Reciprocal Influence Model of Social Power: Emerging Principles and Lines of Inquiry. *Advances in Experimental Social Psychology*, 40(07), 151-192. doi:10.1016/S0065-2601(07)00003-2

Keltner, D. (2008). A reciprocal influence model of social power: Emerging principles and lines of inquiry. In Van Kleef,Gerben A., Chen,Serena, & Kraus,Michael W. (Eds.), (Vol. 40, pp. 151-151). Elsevier.

Kennedy, H. P. (2004). Enhancing Delphi research: methods and results. *Journal of Advanced Nursing*, *45*(5), 504-511. doi:10.1046/j.1365-2648.2003.02933.x

Kim, H., Yoon, J. W., & Crowcroft, J. (2012). Network analysis of temporal trends in scholarly research productivity. *Journal of Informetrics*, *6*(1), 97-110. Elsevier Ltd. doi:10.1016/j.joi.2011.05.006

Klamma, R., Phnam, M. C., & Cao, Y. (2009). You Never Walk Alone: Recommending Academic Events Based on Social Network Analysis. *Complex Sciences*, *4*(1), 657-670.

Klein, K. J., & Sorra, J. S. (1996). The Challenge of Innovation Implementation. *Academy of Management Review*, *21*(4), 1055-1080.
Klusch, M., & Gerber, A. (2002). Dynamic coalition formation among rational agents. *IEEE Intelligent Systems*, *17*(3), 42-47.
Kohlberg, E. (1971). On the nucleolus of a characteristic function game. *SIAM Journal on Applied Mathematics, 20*(1), 62-66.
Kogut, B. (1989). The stability of joint ventures: Reciprocity and competitive rivalry. *The Journal of Industrial Economics*, *38*(2), 183-198.
Kotter, J. P. (1996). *Leading Change* (p. 208). Harvard Business Press. Retrieved from http://books.google.nl/books?id=ib9Xzb5eFGQC&lpg=PR11&ots=Mvn9eIASZ_&dq=kotter 1996&lr&pg=PR11#v=onepage&q=kotter 1996&f=false
Kraatz, M. S. (1998). Learning by Association? Interorganizational Networks and Adaptation to Environmental Change. *The Academy of Management Journal*, *41*(6), 621. doi:10.2307/256961
Krackhardt, D. (1990). Assessing the Political Landscape: Structure, Cognition, and Power in Organizations. *Administrative Science Quarterly*, *35*(2), 342-369.
Kratzer, J., & Lettl, C. (2008). A Social Network Perspective of Lead Users and Creativity: An Empirical Study among Children. *Creativity and Innovation Management*, *17*(1), 26-36. doi:10.1111/j.1467-8691.2008.00466.x
Lagoze, C., & Van de Sompel, H. (2001). The Open Archives Initiative: Building a low-barrier interoperability framework. *Russian Digital Libraries Journal*, *4*(6). Retrieved from http://www.elbib.ru/index.phtml?page=elbib/eng/journal/2001/part6/LS
Lambiotte, R., & Panzarasa, P. (2009). Communities, Knowledge Creation, and Information Diffusion. *Journal of Informetrics*, *3*(3), 180-190. Elsevier.
Lankau, M. J., & Scandura, T. A. (2002). An Investigation of Personal Learning in Mentoring Relationships: Content, Antecedents, and Consequences. *Academy of Management Journal*, *45*(4), 779-790.
La Porta, R., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. W. (1997). Trust in large organizations. *AEA* Papers *and Proceedings*, *87*(2), 310-321.
Latané, B., Williams, K., & Harkins, S. (1979). Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology*, *37*(6), 822-832.
Lave, J. (1991). *Situating Learning in Communities of Practice*.
Lazarsfeld, P. F., & Merton, R. K. (1954). Friendship as a social process: A substantive and methodological analysis. In M. Berger, T. Abel, & C. H. Page (Eds.), *Freedom and Control in Modern Society* (Vol. 18, pp. 18-66). Van Nostrand. Retrieved from http://www.questia.com/PM.qst?a=o&docId=23415760
LeBoeuf, R. a., & Shafir, E. (2003). Deep thoughts and shallow frames: on the susceptibility to framing effects. *Journal of Behavioral Decision Making*, 16(2), 77-92. doi:10.1002/bdm.433

Leydesdorff, L., & Wagner, C. S. (2008). International collaboration in science and the formation of a core group. *Journal of Informetrics*, *2*(4), 317-325. doi:10.1016/j.joi.2008.07.003

Liden, R. C., Wayne, S. J., Jaworski, R. A., & Bennett, N. (2004). Social loafing: A field investigation. *Journal of Management*, *30*(2).

Linstone, H. A., & Turoff, M. (1975). *The Delphi method: Techniques and applications*. Addison-Wesley Pub. Co., Advanced Book Program. Retrieved from http://www.is.njit.edu/pubs/delphibook/

Linton, J. D., Tierney, R., & Walsh, S. T. (2011). Publish or Perish: How Are Research and Reputation Related? *Serials Review*, *37*(4), 244-257. Elsevier Inc. doi:10.1016/j.serrev.2011.09.001

Ma, T., & Nakamori, Y. (2005). Agent-based modeling on technological innovation as an evolutionary process. *European Journal of Operational Research*, *166*(3), 741-755. doi:10.1016/j.ejor.2004.01.055

Maurer, U., Smailagic, A., Siewiorek, D. P., & Deisher, M. (2006). Activity Recognition and Monitoring Using Multiple Sensors on Different Body Positions. *International Workshop on Wearable and Implantable Body Sensor Networks (BSN'06)*, 113-116. Ieee. doi:10.1109/BSN.2006.6

McCarter, M. W., Rockmann, K. W., & Northcraft, G. B. (2009). Is it even worth it? The effect of loss prospects in the outcome distribution of a public goods dilemma. *Organizational Behavior and Human Decision Processes*.

McKenna, H. P. (1994). The Delphi technique: a worthwhile research approach for nursing? *Journal of Advanced Nursing*, *19*(6), 1221-1225. doi:10.1111/j.1365-2648.1994.tb01207.x

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, *27*(1), 415-444. doi:10.1146/annurev.soc.27.1.415

Merton, Robert K. (1984). The Focussed Interview and Focus Groups. *Public Opinion Quarterly, 51*, 550-566.

Mika, P. (2006). *Social Networks and the Semantic Web*. Retrieved from http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1410816

Milkovich, G. T., Annoni, A. J., & Mahoney, T. A. (1972). The Use of the Delphi Procedures in Manpower Forecasting. *Management Science*, *19*(4), 381-388.

Moscovici,S., & Zavalloni,M. (1969). The Group as a Polarizer of Attitudes. *J Personality Soc Psychol*.

Nguyen, L., & Shanks, G. (2009). A framework for understanding creativity in requirements engineering. *Information and Software Technology*, *51*(3), 655-662. Elsevier B.V. doi:10.1016/j.infsof.2008.09.002

Nijstad, B. A., Stroebe, W., & Lodewijkx, H. F. M. (2003). Production blocking and idea generation: Does blocking interfere with cognitive processes? *Journal of Experimental Social Psychology*, *39*(6), 531-548.

Nowak, M. a, & Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*, *437*(7063), 1291-8. doi:10.1038/nature04131

Olson, J. S., Grudin, J., & Horvitz, E. (2004). Toward Understanding Preferences for Sharing and Privacy. *Public Policy*.

Osborn, A. F. (1954). *Applied Imagination*. New York, NY, USA: Scribner.
O'Neill, S. J., Osborn, T. J., Hulme, M., Lorenzoni, I., & Watkinson, A. R. (2008). Using expert knowledge to assess uncertainties in future polar bear populations under climate change. *Journal of Applied Ecology*, *45*(6), 1649-1659. doi:10.1111/j.1365-2664.2008.01552.x
Owen-Smith, J., Powell, W.W. (2004). Knowledge Networks as Channels and Conduits: The Effects of Spillovers in the Boston Biotechnology Community, *Organization Science*, *15*(1), p.5-21, January 2004 doi>10.1287/orsc.1030.0054
Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*, 2(1-2), 1-135.
Pazzani, M. J. (1999). A framework for collaborative, content-based and demographic filtering. *Artificial Intelligence Review*, *13*(5), 393–408. Springer. Retrieved from http://www.springerlink.com/index/T61V33QX45465531.pdf
Penuel W. R., Riel M., Krause A.E. & Frank K.A. (2009), Analyzing Teachers' Professional Interactions in a School as Social Capital: A Social Network Approach, *Teachers College Record Volume 111*, Number 1, January 2009, pp. 124–163. http://www.tcrecord.org/content.asp?contentid=15174
Perry-Smith, J. (2006). Social Yet Creative: The Role of Social Relationships in Facilitating Individual Creativity. *Academy of Management Journal*, *49*(1), 85-101.
Phelan, S. (2002). Cognitive capacity as competitive advantage: a simulation test. *Simulation Modelling Practice and Theory*, *10*(5-7), 455-471. doi:10.1016/S1569-190X(02)00121-1
Pifer M. (2011), *Such a Dirty Word: Networks and Networking in Academic Departments*, PhD Dissertation, The Pennsylvania State University. DAI-A 71/09, p.289, Mar 2011
Pinto, J. K., & Kharbanda, O. P. (1996). How to fail in project management (without really trying). *Business Horizons*, *39*(4), 45-53. doi:10.1016/S0007-6813(96)90051-8
Pirolli, P. (2009). An Elementary Social Information Foraging Model. *Proceedings of the 27th international conference on Human factors in computing systems (CHI'09)*, pp. 605-614. Boston, MA, USA.
Prud'hommeaux, E., & Seaborne, A. (2012). SPARQL Query Language for RDF. *W3C*. Retrieved March 8, 2012, from http://www.w3.org/TR/rdf-sparql-query/
Rajagopal, K., Joosten-ten Brinke, D., Van Bruggen, J., & Sloep, P.B. (2012). Understanding Personal Learning Networks: their structure, content and the networking skills needed. *First Monday*, *17*(1), 1-29.
Ratner, R. K., & Miller, D. T. (2001). The norm of self-interest and its effects on social action. *Journal of personality and social psychology*, *81*(1), 5-16. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11474725
Reinhardt, W. (2012). *Awareness Support for Knowledge Workers in Research Networks*. Open Universiteit. Retrieved from http://dspace.ou.nl/handle/1820/4193
Reinhardt, W., Meier, C., Drachsler, H., & Sloep, P. (2011). Analyzing 5 years of EC-TEL proceedings. In C. D. Kloos, D. Gillet, R. M. C. Garcìa, F. Wild, & M. Wolpers

(Eds.), *Towards Ubiquitous Learning. Proceedings of the 6th European Conference on Technology Enhanced Learning*.

Reinhardt, W., Mletzko, C., Drachsler, H., & Sloep, P. (2011). AWESOME: A widget-based dashboard for awareness-support in Research Networks. *Proceedings of the 2nd PLE Conference*.

Rice, K. (2009). Priorities in K – 12 Distance Education: A Delphi Study Examining Multiple Perspectives on Policy , Practice , and Research. *Educational Technology & Society, 12*, 163-177.

Ring, P.S., & Van de Ven, A.H. (1994) Developmental processes of cooperative interorganizational relationships, *Academy of Management Review*, *19*(1), 90-118.

Rogers, Everett M. (1995) *Diffusion of innovations*, 4th ed., New York: Free Press.

Rokkan, A. I., & Buvik, A. (2003). Inter-firm cooperation and the problem of free riding behavior: an empirical study of voluntary retail chains. *Journal of Purchasing and Supply Management, 9*(5-6), 247-256. doi:10.1016/j.pursup.2003.09.004

Rose, J. D. (2011). Diverse Perspectives on the Groupthink Theory – A Literary Review. *Emerging Leadership Journeys, 4*(1), 37-57.

Ruch, J., Heinrich, L., Bilde, T., & Schneider, J. M. (2009). Relatedness facilitates cooperation in the subsocial spider, Stegodyphus tentoriicola. *BMC evolutionary biology*, *9*, 257. doi:10.1186/1471-2148-9-257

Rusman, E., Bruggen, J. V., Cörvers, R., Sloep, P., & Koper, R. (2009). From pattern to practice: evaluation of a design pattern fostering trust in Virtual teams. *Computers in Human Behavior*, *25*(5), 1010-1019.

Rusman, E., Bruggen, J. V., Sloep, P., & Koper, R. (2009). Fostering trust in virtual project teams: Towards a design framework grounded in a TrustWorthiness ANtecedents(TWAN) schema. *International Journal of Human-Computer Studies*, *68*(11), 834-850.

Ruthledge, R. W. (2011). Escalation of Commitment in Groups and the Moderating Effects of Information Framing. *Journal of Applied Business Research*, *11*(2).

Schank, R. C. (1995). What We Learn When We Learn by Doing (Technical Report No. 60). *Sciences-New York* (pp. 1-37).

Schmeidler, D. (1969). The nucleolus of a characteristic function game. *SIAM Journal on Applied Mathematics*, 1163-1170.

Schmid, K. (1996). Making AI systems more creative: the IPC-model Processes / Operators. *Knowledge-Based Systems*, 9, 385-397.

Schreiner, M., Kale, P., & Corsten, D. (2009). WHAT REALLY IS ALLIANCE MANAGEMENT CAPABILITY AND HOW DOES IT IMPACT ALLIANCE OUTCOMES AND SUCCESS? *Strategic Management Journal*, *30*, 1395-1419.

Schwarz, N., & Ernst, A. (2009). Agent-based modeling of the diffusion of environmental innovations–An empirical approach. *Technological forecasting and social change*, *76*(4), 497–511. Elsevier. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0040162508000875

Seah, C., Sierhuis, M., & Clancey, W. J. (2005). Multi-agent Modeling and Simulation Approach for Design and Analysis of MER Mission Operations. *International*

Conference on Human-Computer Interface Advances for Modeling and Simulation (SIMCHI'05).

Selten, R. (1998). Features of Experimentally Observed Bounded Rationality. *European Economic Review*, *42*, 413-436.

Shapley, L. S. (1953). A value for n-person games. Contribution to the Theory of Games. (H. W. Kuhn & A. W. Tucker, Eds.)*Annals of Mathematics Studies*, *2*, 28.

Shubik, M. (1971). The Dollar Auction game: a paradox in noncooperative behavior and escalation. *Journal of Conflict Resolution*, *15*(1), 109-111.

Sie, Rory L.L., Bitter-Rijpkema, M., & Sloep, P. B. (2010a). Coalition Formation in Networked Innovation: Directions for Future Research. In L. Dirckinck-Holmfeld, V. Hodgson, C. Jones, M. de Laat, D. McConnell, & T. Ryberg (Eds.), *Proceedings of the Seventh International Conference on Networked Learning 2010*. Aalborg, Denmark. Retrieved from http://hdl.handle.net/1820/2173

Sie, Rory L L, Bitter-Rijpkema, M., & Sloep, P. B. (2010b). A Simulation for Content-based and Utility-based Recommendation of Candidate Coalitions in Virtual Creativity Teams. *Procedia Computer Science: RecSysTEL2010*, *1*(2), 2883-2888.

Sie, R. L. L., Bitter-Rijpkema, M., & Sloep, P. B. (2011). What's in it for me? Recommendation of Peers in Networked Innovation. *Journal of Universal Computer Science, 17*(12), 1659-1672.

Sie, R.L.L., Berlanga, A.J., Rajagopal, K., Pannekeet, K., Drachsler, H., Fazeli, S. Sloep, P.B. (2012). Social Tools for Networked Learning: Current and Future Research Directions. *Proceedings of the Eighth International Conference on Networked Learning 2012*. Maastricht, The Netherlands

Sie R. L. L., Drachsler, H., Bitter-Rijpkema, M., & Sloep, P. B. (2012). To whom and why should I connect? Co-author Recommendation based on Powerful and Similar Peers. *International Journal of Technology Enhanced Learning, 4*(1), 121-137, DOI: 10.1504/IJTEL.2012.048314

Sie, R.L.L., Pataraia, N., Boursinou, E., Rajagopal, K., Falconer, I., Margaryan, A., Bitter-Rijpkema, M., Littlejohn, A., Sloep, P.B. (submitted). Goals, Motivation for, and Outcomes of Learning through Networks: Results of a Tweetstorm

Sie, R.L.L., Ullmann, T. D., Rajagopal, K., Cela, K. (submitted). The Technology-enhanced Learning Christmas Carol: The Ghosts from the Past, Present, and Future for Social Network Analysis.

Sie, R.L.L., Van Engelen, B.J., Bitter-Rijpkema, M., & Sloep, P.B. (submitted). COCOON CORE: CO-author Recommendations based on Betweenness Centrality and Interest Similarity

Sie R. L. L., Bitter-Rijpkema, M., & Sloep, P. B. (submitted). If We Work Together, I Will Have Greater Power: Coalitions in Networked Innovation.

Siemens, G. (2005). Connectivism: A Learning Theory for the Digital Age. *International Journal of Instructional Technology and Distance Learning*, *2*(1), 3-10. elearnspace. Retrieved from http://www.elearnspace.org/Articles/connectivism.htm

Simon, H. A. (1982). *Models of bounded rationality*. MIT Press.

Simon, H. A. (1991). Bounded rationality and organizational learning. *Organization Science*, *2*(1), 125-134.

Sinha, R., & Swearingen, K. (2002). The role of transparency in recommender systems. *CHI ’02 extended abstracts on Human factors in computing systems - CHI '02*, 830. New York, New York, USA: ACM Press. doi:10.1145/506443.506619

Sloep, P., & Berlanga, A. J. (2011). Learning Networks, Networked Learning [Redes de Aprendizaje, Aprendizaje en Red]. *Comunicar*, *XIX*(37), 55-63. Retrieved from http://dx.doi.org/10.3916/C37-2011-02-05

Sloep, P. B., Van der Klink, M., Brouns, F., Van Bruggen, J., & Didderen, W. (Eds.) (2011). *Leernetwerken; Kennisdeling, kennisontwikkeling en de leerprocessen*. Houten, Nederland: Bohn, Stafleu, Van Loghum.

Song, F. (2009). Intergroup trust and reciprocity in strategic interactions: Effects of group decision-making mechanisms. *Organizational Behavior and Human Decision Processes*, *108*(1), 164-173. doi:10.1016/j.obhdp.2008.06.005

Stahl, G. (2005). Group cognition in computer-assisted collaborative learning. *Journal of Computer Assisted Learning*, *21*(2), 79-90. doi:10.1111/j.1365-2729.2005.00115.x

Swan, K. (2002). Building Learning Communities in Online Courses : the importance of interaction. *Education, Communication & Information*, *2*(1), 23-49. doi:10.1080/1463631022000005016

Swan, J., & Scarbrough, H. (2005). The politics of networked innovation. *Human Relations*, *58*(7), 913-943. doi:10.1177/0018726705057811

Tetlock, P. E. (1992). The impact of accountability on judgment and choice: Toward a social contingency model. *Advances in Experimental Social Psychology Volume*, *25*.

Tsai, W., & Ghoshal, S. (1998). Social Capital and Value Creation : The Role of Intrafirm Networks. *The Academy of Management Journal*, *41*(4), 464-476.

Turner, C. W., Lewis, J. R., & Nielsen, J. (2006). Determining Usability Test Sample Size. *International Encylcopedia of Ergonomics and Human Factors*, *3*(2), 3084-3088.

Ullmann, T. D., Wild, F., Scott, P., Duval, E., Parra, G., Reinhardt, W., Heinze, N., et al. (2010). Components of a Research 2.0 Infrastructure. In V. Dimitrova, M. Wolpers, & P. Kirschner (Eds.), *Proceedings of the 5th European Conference on Technology-Enhanced Learning (ECTEL’10)* (pp. 590-595). Barcelona, Spain: Springer.

Väljataga, T., & Fiedler, S. (2009). Supporting students to self-direct intentional learning projects with social media. *Educational Technology & Society*, *12*(3), 58-69.

Van Rosmalen, P., Sloep, P., Kester, L., Brouns, F., de Croock, M., Pannekeet, K., & Koper, R. (2007). A learner support model based on peer tutor selection. *Journal of Computer Assisted Learning*, 070528071048002-??? doi:10.1111/j.1365-2729.2007.00245.x

Vassileva, J., McCalla, G., & Greer, J. (2003). Multi-agent multi-user modeling in I-Help. *User Modeling and User-Adapted Interaction*, *13*(1), 179–210. Springer. Retrieved from http://www.springerlink.com/index/j542025743037675.pdf

Virzi, R. A. (1992). Refining the Test Phase of Usability Evaluation : How Many Subjects Is Enough? *Human Factors*, *34*(4), 457-468.

Von Neumann, J., & Morgenstern, O. (1945). Theory of games and economic behavior. *Bull. Amer. Math. Soc. 51 (1945),* 498-504. DOI: 10.1090/S0002-9904-1945-08391-8 PII: S, 2(9904), 08391-8.

Vygotsky, L. (1978). *Mind in society*. Cambridge, Massachusetts: Harvard University Press.

Wallas, G. (1976). Stages in the creative process. *The creativity question*, 69-73.

Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications* (p. 825). Cambridge: Cambridge University Press.

Weebl's Stuff. (2008). The mango song. Retrieved December 9, 2011, from http://www.weebls-stuff.com/songs/mango/

Wenger, E. (1999). *Communities of practice: learning, meaning, and identity* (p. 318). Cambridge University Press. Retrieved from http://books.google.com/books?hl=en&lr=&id=heBZpgYUKdAC&pgis=1

Wesch, M. (2009). From Knowledgable to Knowledge-able: Learning in New Media Environments. *Academic Commons*, (January).

West, M. A. (2002). Sparkling Fountains or Stagnant Ponds: An Integrative Model of Creativity and Innovation Implementation in Work Groups. *Applied Psychology: An International Review*, *51*(3), 355-424.

Whitworth, B., & Whitworth, A. P. (2010). The social environment model: Small heroes and the evolution of human society. *First Monday*, *15*(11), 1-18. Retrieved from http://www.firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3173

Winter, S. G. (2000). The satisficing principle in capability learning. *Strategic Management Journal*, 981-996.

Wooldridge, M. (1998). Agent-Based Computing. *Interoperable Communication Networks*, *1*(1), 71-97.

Wu, D. D., Kefan, X., Hua, L., Shi, Z., & Olson, D. L. (2010). Modeling technological innovation risks of an entrepreneurial team using system dynamics: An agent-based perspective. *Technological Forecasting and Social Change*, *77*(6), 857-869. Elsevier. doi:doi:10.1016/j.techfore.2010.01.015

Yazici, H. J. (2005). A study of collaborative learning style and team learning performance. *Education and Training*, *47*(3), 216-229. doi:10.1108/00400910510592257

Zimmerman, B. J., Bandura, A., & Martinez-pons, M. (1992). Self-Motivation for Academic Attainment: The Role of Self-Efficacy Beliefs and Personal Goal Setting. *American Educational Research Journal*, *29*(3), 663-676.

Appendices

Appendix A - Statements per cluster at the level of seven core clusters

cluster	name	statements
1	Sharing	a1, a2, a3, a4, a5
2	Motivation	a6, a7, a8, a9, a10, a11, a12, a13, a14, a15, a16, a17, a18, a28, a29, a30, a31, a32, a33, a34, a35, a36, a70, a71, a72, a73, a74, a75, a76, a77, a78, a79
3	Perceived value of the network	a19, a20, a25, a26, a27, a40, a41, a42, a43, a44, a45, a46, a47, a48, a49, a54
4	Feedback	a21, a22, a23, a24
5	Personal learning	a37, a38, a39, a55, a56, a57, a58, a80, a81, a82, a83
6	Trust and support	a50, a51, a52, a53, a59, a60, a61, a62, a63
7	Peer characteristics and value	a64, a65, a66, a67, a68, a69

Appendix B - Questions regarding quality of recommendations

1a. Individual Recommendation: How do you value the recommendation that is generated if you control the sliders yourself?
1b. Individual Recommendation: How do you value the recommendation that is generated if the slider for influence is set to 100?
1c. Individual Recommendation: How do you value the recommendation that is generated if the slider for interest similarity is set to 100?
2a. Default User Recommendation: How do you value the recommendation that is generated if you control the sliders yourself?
2b. Default User Recommendation: How do you value the recommendation that is generated if the slider for influence is set to 100?
2c. Default User Recommendation: How do you value the recommendation that is generated if the slider for interest similarity is set to 100?

Appendix C: SUS questionnaire

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.

Summary

Summary

The central question of this thesis, is:

How can we assemble individuals that want to cooperate to create something new?

The perspective that this thesis' research takes as a starting point is cooperation networks. Cooperation networks are social networks of individuals that have the intention to go into the same direction, though they do not necessarily have the same goal. Inherently, this distinguishes cooperation from collaboration, in which individuals do share a common goal.

A nice example of a cooperation network is the Automobile Manufacturers Association. The origin of this association lies in the dispute that George Selden and Henry Ford had in the early 1900s. George Selden had patented a 'road engine', a car-like vehicle, and started collecting money from other car manufacturers. Henry Ford refused to pay Selden, arguing that the road engine could not work. Selden took Ford to court, and eventually, the judge decided that Selden had to build and test the road engine. Indeed, the road engine did not work. Ford won the case and decided to found the Automobile Manufacturers Association to openly share patents among car manufacturers. In other words, they formed a network of car manufacturers by having the common intention to share their patents. They did however have their distinct goals of making money for themselves and staying ahead of the competition.

The example shows how cooperation in practice can take place. It is, however, easier to state that you intend to cooperate than to actually do it. Individuals generally encounter four types of problems when they want to cooperate (Figure 9.1; *Chapter 1*). Firstly, they are hampered by intrapersonal problems, such as bounded rationality, framing and information overload. Secondly, they are prone to interpersonal problems, such as self-interest, social loafing, and lack of trust. Thirdly, they face procedural and structural problems, such as deciding which stage in innovation (e.g. problem identification, idea generation, idea implementation) calls for a homogeneous rather than a heterogeneous group of cooperating people. Finally, people experience exogenous problems, such as a firm's culture, or a lack of funding.

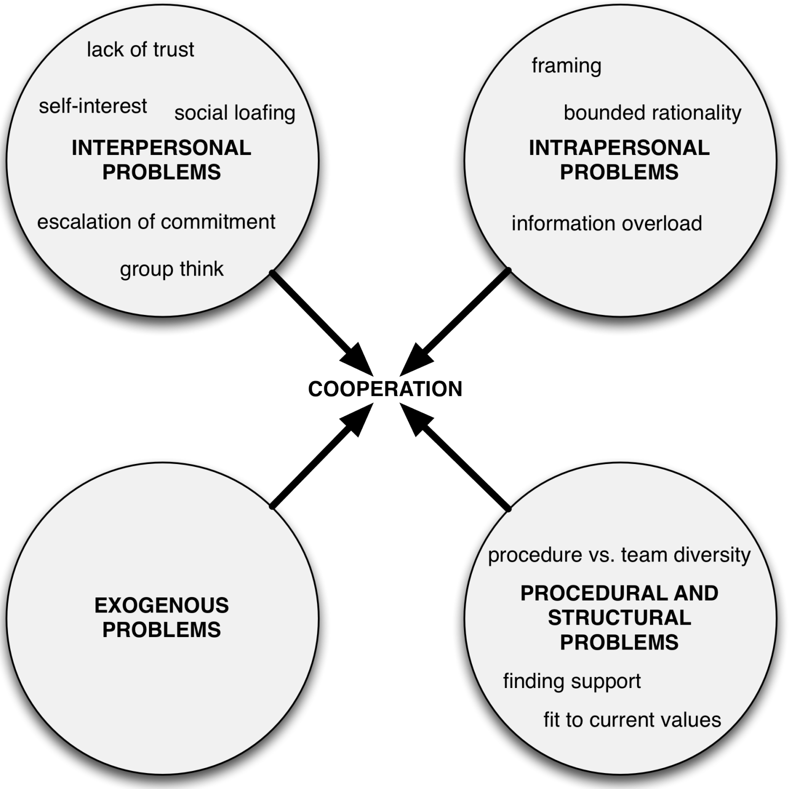


Figure 9.1 Four main types of problems in cooperation networks.

Cooperating individuals encounter a myriad of problems, and even for a thesis, this is too large a number to crack. We therefore restricted ourselves to studying how to solve interpersonal and intrapersonal problems. To solve these problems (as will be shown, mainly through carrying out interventions), it is necessary to have a thorough understanding of the factors that play a role in cooperation networks, and of the way they interact with one another. Each chapter in this thesis deals with different aspects of the main research question, broken down in research subquestions (Figure 9.2).

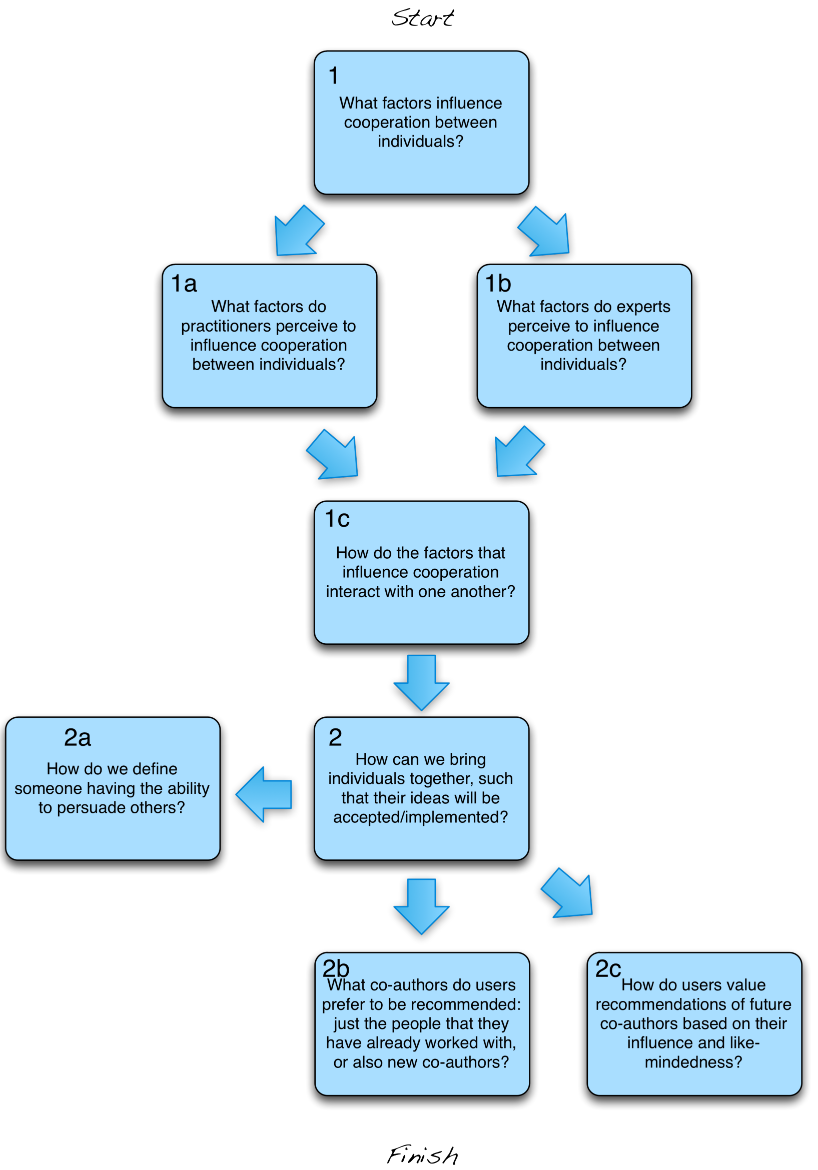


Figure 9.2. Main structure of the thesis.

Contributions to theory

First, we investigated which factors practitioners of a special type of cooperation networks, learning networks, perceive to influence their personal, professional learning (*Chapter 2*). We employed a new method to identify these factors, called the *Tweetstorm*, which is an amalgamation of tweets (microblog messages via

Twitter) and the brainstorm technique to generate ideas. After aggregation of the statements that were in the tweets, we asked experts to categorize the statements to arrive at a set of core clusters of factors. The results show seven core clusters of factors, and fourteen subclusters that practitioners perceive to drive their personal learning (Figure 9.3).

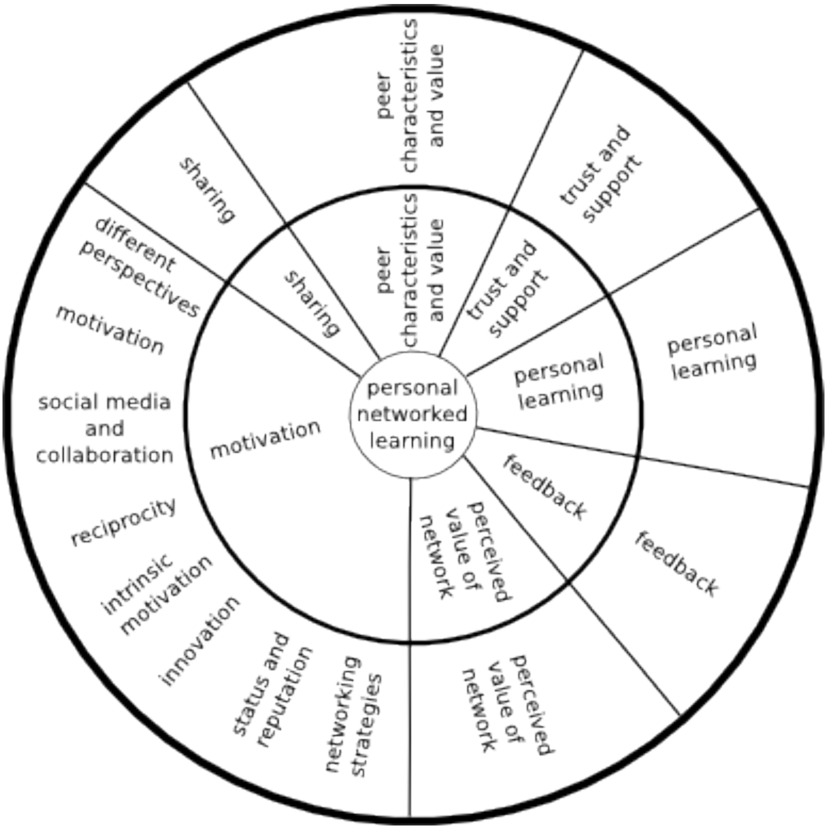


Figure 9.3. Core clusters of factors that influence personal learning, as perceived by personal, professional networked learners.

The ensemble of factors that practitioners identify as influencing their personal professional learning in networks does not cover each and every factor that actually influences cooperation in networks. Intensely studying the available literature covers another part of these factors, but to make sure we did not miss out on any factors or recent developments, we asked senior experts to identify the factors that influence cooperation networks from their perspective (*Chapter 3*). We employed an online version of the Delphi method for discussion and consensus finding among experts called *eDelphi* to help the experts identify the factors.

The eDelphi brought forward four core clusters of factors that influence cooperation:

1. Personal characteristics
2. Diversity
3. Effective cooperation
4. Managerial aspects

The experts were based in various disciplines. This, and the fact that they were mainly seniors, in contrast to the practitioners in Chapter 2, probably led to a more high-level view of factors that influence cooperation in networks.

Simulation

The factors that resulted from the literature study and the experiments in Chapters 2 and 3 formed the basis for a first simulation model (*Chapter 4*). The simulation model tries to capture the interplay of factors in innovation networks, another kind of cooperation networks. The results showed that, to have their idea implemented, individuals with low power can loaf on individuals that have higher power. This provided an interesting point of view for, for instance, research cooperation. If you manage to convince an individual with higher power of your research idea, you may have a higher chance of being accepted by the research field's community.

We then took the simulation model from Chapter 4 as a starting point to further investigate the interplay of factors (*Chapter 5)*. We made use of the so-called *parameter sweeping* method, which entails varying all factors within a predefined range during a series of simulation runs. This particularly allows one to study the subtle behaviour of the model. The results showed that a good position of an individual in the network, a so-called high betweenness centrality, is predictive of the average power of a winning coalition between individuals. Particularly, as the average betweenness of the individuals in a winning coalition increases, its average power decreases. This means that when you have high betweenness as an individual, it is easier to stand out as a coalition and have success implementing your idea.

Researcher support

As a way to view these findings from a practical angle, we focused on intervening in the practice of doing research. Every researcher has good ideas, but not all good ideas always find their way to a publication in a journal. Researchers are in need of strategic partnerships to increase their outreach, but are at the same in need of finding peers who are performing research on the same subject. The COCOON system that we developed assists researchers by recommending them key individuals (*Chapter 6*).

The research network that we analysed to generate recommendations is extracted from an institution's local DSpace repository that contains publications and their metadata. For each article in that database we extracted its co-authors, which form a co-author relationship in a co-authorship network, and its keywords. The system used betweenness centrality to identify powerful peers in the co-authorship network. To identify like-minded co-authors, the system used the similarity between the keywords that authors use to describe their documents.

The recommendation algorithm takes the weighted average of both the betweenness and the keyword similarity of co-authors to the target user. Users were presented two lists: one with merely new co-authors, and one with new and existing co-authors. The results showed that users prefer to have existing co-authors recommended as well, because they are relatively unfamiliar with the work of co-authors that they did not yet work with.

The COCOON system was succeeded by COCOON CORE. And like COCOON, it provides a means to take a practical look at cooperation (*Chapter 7*). COCOON CORE focuses on further empowerment of the user by giving them the opportunity to adjust the balance between finding powerful peers and like-minded peers themselves. Also, it presents author pages to give further insight into what authors write, what their output quality is, and how authors are related to one another. Finally, it presents keyword pages that show their quality and how they are related to one another.

We conducted an evaluation experiment among researchers of the institution that hosts the DSpace database and the results showed that the participants value the ability to modify the recommendation algorithm themselves. In general, the recommendations were scored moderately positively. COCOON CORE was also tested for its user friendliness. It scored moderately positively on usability, and particularly its learnability scores were high. Future work on COCOON CORE should focus more on the integration of its services.

Conclusion

This thesis focused on interpersonal and intrapersonal problems in cooperation networks. We specifically aimed at overcoming bounded rationality by aiding the decision process of an individual in search of new cooperation in her network. Also, we aimed at decreasing the information overload that individuals typically encounter when they search their network for valuable peers. Each chapter in this thesis aimed at solving a specific subproblem that one comes across in the step-by-step process to successfully introduce a system that assists cooperation in networks.

The results show that a system that recommends powerful and like-minded peers for cooperation is valued among users and thus has potential. In this era of social media, it may be particularly interesting to pursue further research in the direction of network-based recommender systems. From the perspective of social network research, it is time to take the next step and create a social network theory that

informs interventions instead of resting content with merely analysing social networks.

Samenvatting

Samenvatting

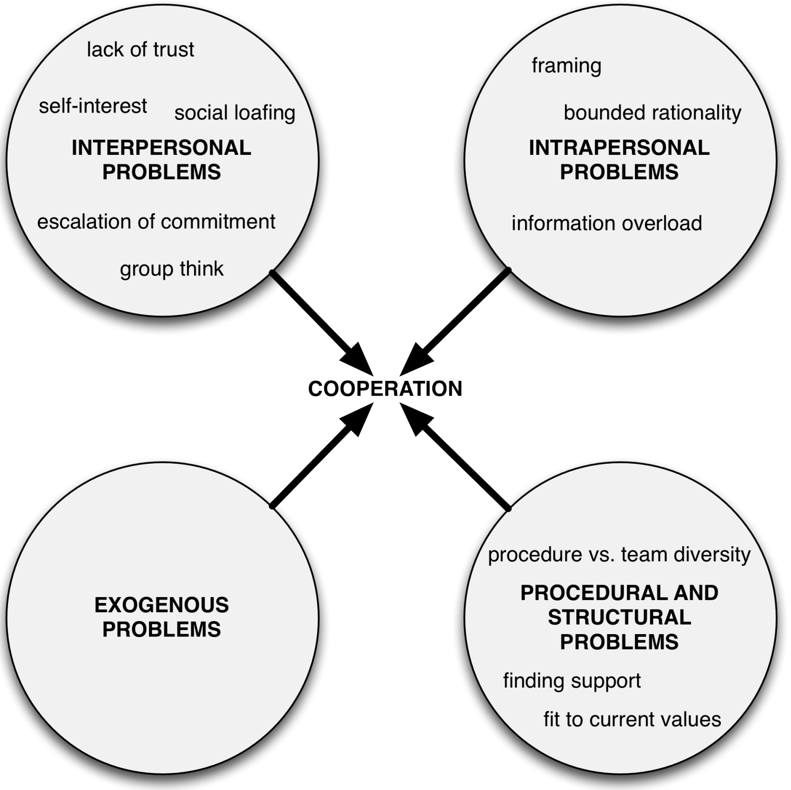
De centrale vraag van dit proefschrift is:

Hoe kunnen we individuen samenbrengen die willen samenwerken om iets nieuws te creëren?

Het perspectief dat het onderzoek in dit proefschrift als uitgangspunt neemt is coöperatienetwerken. Coöperatienetwerken zijn sociale netwerken van individuen die de intentie hebben om dezelfde richting op te gaan, hoewel ze niet per se hetzelfde doel voor ogen hebben. Dit onderscheidt coöperatie van collaboratie, waarin individuen een gemeenschappelijk doel hebben.

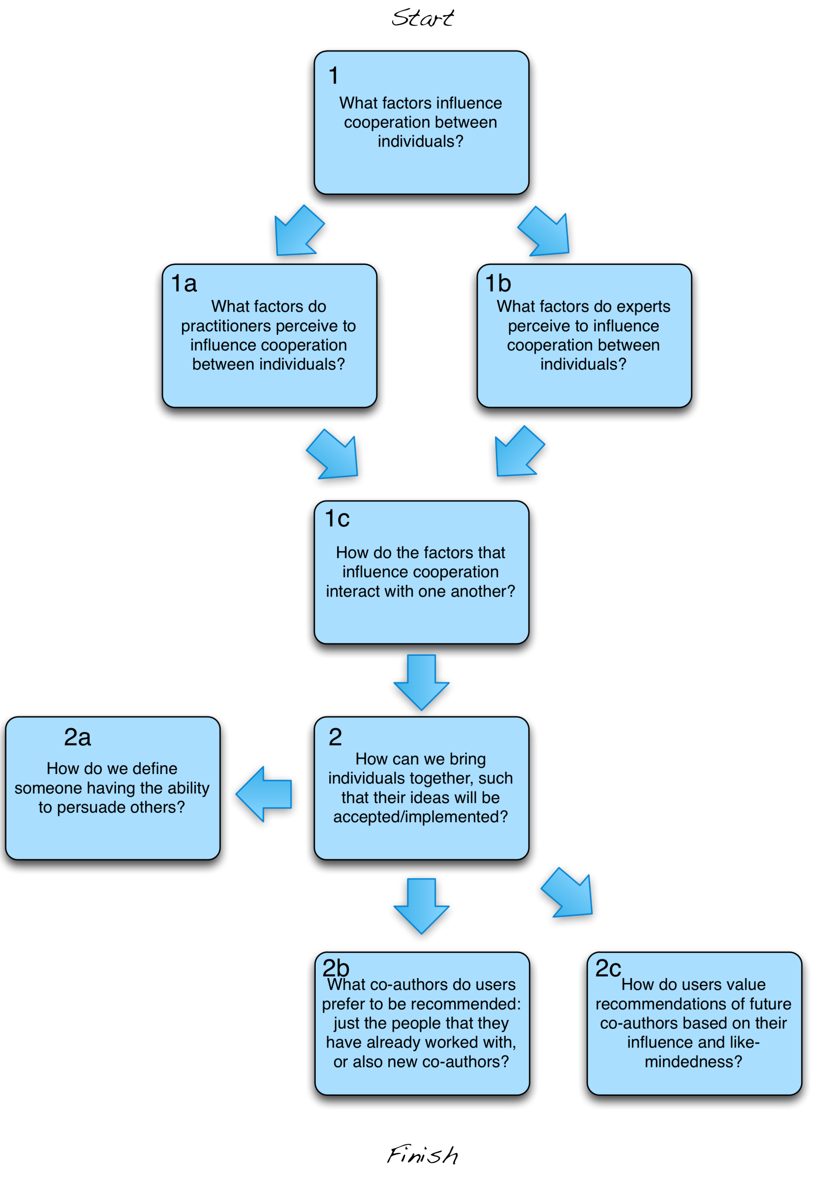
Een mooi voorbeeld van een coöperatienetwerk is de Automobile Manufacturers Association. De oorsprong van deze associatie ligt in het geschil dat George Selden en Henry Ford hadden in de vroege 20^{e} eeuw. George Selden had een patent op een 'road-engine', een auto-achtige voertuig, en begon met het inzamelen van geld van andere autofabrikanten. Henry Ford weigerde Selden te betalen, met het argument dat de road engine niet zou kunnen werken. Selden daagde Ford voor de rechter, en uiteindelijk besloot de rechter dat Selden de road engine moest bouwen en testen. Uiteraard heeft de road engine nooit gewerkt. Ford won de zaak en besloot de Automobile Manufacturers Association op te richten om zodoende openlijk patenten te delen met autofabrikanten. Met andere woorden, ze vormden een netwerk van autofabrikanten door de gemeenschappelijke intentie om hun patenten te delen. Ze hadden echter wel hun eigen doelen om zelf winst te maken en de concurrentie voor te blijven.

Dit voorbeeld laat zien hoe coöperatie in de praktijk kan plaatsvinden. Aangeven dat je van plan bent om samen te werken is makkelijker gezegd dan gedaan. Mensen komen over het algemeen vier soorten problemen tegen wanneer ze willen samenwerken (Figuur 9.1; hoofdstuk 1). Ten eerste worden ze gehinderd door intrapersoonlijke problemen, zoals begrensde rationaliteit, *framing* en informatieoverdaad. Ten tweede zijn ze gevoelig voor interpersoonlijke problemen, zoals eigenbelang, meeliften, en gebrek aan onderling vertrouwen. Ten derde worden zij geconfronteerd met procedurele en structurele problemen, zoals beslissen welke stap in innovatie (bv. probleemidentificatie, het genereren van ideeën, implementatie van ideeën) vraagt om een homogene in plaats van een heterogene groep samenwerkende mensen. Tot slot ervaren mensen exogene problemen, zoals de cultuur van een bedrijf, of een gebrek aan financiering.



Figuur 9.1 Vier belangrijke soorten problemen in samenwerkingsnetwerken.

Samenwerkende individuen worden geconfronteerd met een groot aantal problemen, en zelfs voor een proefschrift is dit een te groot aantal om op te lossen. We hebben ons daarom beperkt tot het bestuderen hoe de interpersoonlijke en intrapersoonlijke problemen op te lossen. Voor het oplossen van deze problemen (zoals vooral zal blijken door het uitvoeren van interventies), is een grondige kennis noodzakelijk van de factoren die een rol spelen bij netwerken voor samenwerking en van de manier waarop ze invloed op elkaar hebben. Elk hoofdstuk in dit proefschrift behandelt verschillende aspecten van de centrale onderzoeksvraag, onderverdeeld in deelvragen (Figuur 9.2).

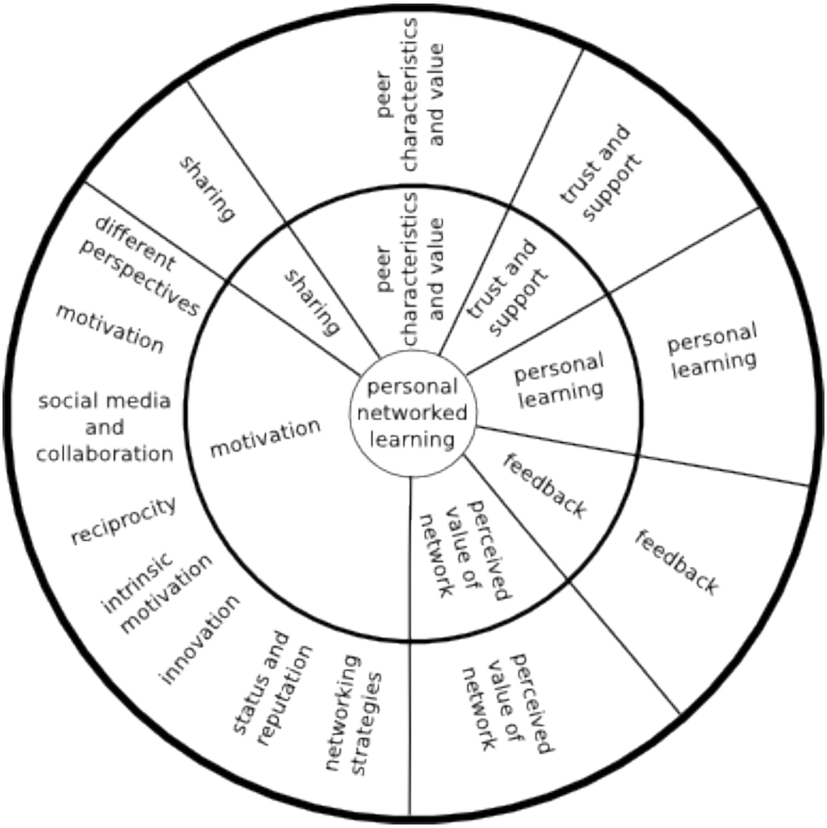


Figuur 9.2. Hoofdstructuur van dit proefschrift.

Bijdragen aan de theorie

Ten eerste hebben we onderzocht welke factoren de deelnemers van een speciaal type van netwerken voor samenwerking, leernetwerken, beschouwen als factoren die hun persoonlijke, professionele leren beïnvloeden (*Hoofdstuk 2*). We gebruikten een nieuwe methode om deze factoren te identificeren, genaamd de

Tweetstorm, wat een samensmelting is van *tweets* (microblog berichten via Twitter) en de brainstormtechniek om ideeën te genereren. Na samenvoeging van de verklaringen die zich in de tweets bevonden vroegen we deskundigen om de verklaringen te categoriseren om tot een set van kernclusters van factoren te komen. De resultaten tonen aan dat de deelnemers zeven fundamentele clusters van factoren en veertien subclusters identificeren die hun persoonlijke leren (figuur 9.3) beïnvloeden.



Figuur 9.3. Kernclusters van factoren die de persoonlijke leren beïnvloeden, zoals waargenomen door persoonlijk, professioneel netwerklerenden.

De factoren die de deelnemers identificeren als van invloed zijnde op hun persoonlijke, professionele leren in netwerken zijn niet direct vertaalbaar naar factoren die van invloed zijn op de samenwerking in netwerken. Het intens bestuderen van de beschikbare literatuur dekt een ander deel van deze factoren af, maar om ervoor te zorgen dat we geen factoren of de recente ontwikkelingen misten, vroegen we senior experts om de factoren die van invloed zijn op netwerken voor samenwerking vanuit hun perspectief (*Hoofdstuk 3*) te identificeren. We gebruikten een online versie van de Delphi-methode voor discussie en het bereiken van consensus onder de experts, genaamd *eDelphi,* om de experts te helpen bij het identificeren van de factoren.

De eDelphi-methode bracht de volgende vier kernclusters van factoren die samenwerking beïnvloeden, naar voren:

1. Persoonlijke kenmerken
2. Verscheidenheid
3. Effectieve samenwerking
4. Leidinggevende aspecten

De experts waren afkomstig uit verschillende disciplines. Dit, en het feit dat ze vrij ervaren waren, in tegenstelling tot de beoefenaars in hoofdstuk 2, heeft waarschijnlijk geleid tot een meer high-level perspectief van de factoren die samenwerking in netwerken beïnvloeden.

Simulatie

De factoren die het resultaat vormden van de literatuurstudie en de experimenten in hoofdstuk 2 en 3 vormden de basis voor een eerste simulatiemodel (*Hoofdstuk 4*). Het simulatiemodel probeerde vast te leggen hoe het samenspel van factoren in innovatienetwerken, een ander soort van samenwerking in netwerken, plaatsvindt. De resultaten toonden aan dat, om hun idee uitgevoerd te krijgen, personen met weinig macht kunnen meeliften op personen die meer macht te hebben. Dit levert een interessant standpunt op voor, bijvoorbeeld, samenwerking in de wetenschap. Als het je lukt om een persoon met meer macht te overtuigen van je onderzoeksidee, heb je een hogere kans om geaccepteerd te worden door de gemeenschap van het onderzoeksveld.

Daarna namen we het simulatiemodel van hoofdstuk 4 als uitgangspunt om verder onderzoek te doen naar het samenspel van factoren (*Hoofdstuk 5*). We hebben gebruik gemaakt van de zogenaamde *parameter sweeping* methode die alle variabelen binnen een vooraf gesteld bereik varieert tijdens een reeks van simulaties. Dit maakt het mogelijk om te bestuderen wat het subtiele gedrag van het model is. De resultaten toonden aan dat een goede positie van een individu in het netwerk, een zogenaamde *high betweenness centrality*, een voorspeller is voor de gemiddelde macht van een winnende coalitie tussen individuen. Vooral als de gemiddelde betweenness centrality van de individuen in een winnende coalitie stijgt, dan daalt de gemiddelde macht. Dit betekent dat wanneer je een hoge betweenness hebt als individu, het makkelijker is om op te vallen als coalitie zijnde en succes te hebben met de acceptatie van je idee.

Ondersteuning voor wetenschappers

Om onze bevindingen te beschouwen vanuit een praktische invalshoek, hebben we ons gericht op het ingrijpen in de praktijk van het verrichten van onderzoek. Iedere wetenschapper heeft goede ideeën, maar niet alle goede ideeën vinden altijd hun weg naar publicatie in een tijdschrift. Wetenschappers hebben behoefte aan strategische samenwerkingsverbanden om hun bereik te vergroten, maar ze zijn tegelijkertijd op zoek naar collega's die onderzoek verrichten naar hetzelfde onderwerp. Het COCOON systeem dat we ontwikkelden helpt wetenschappers door ze waardevolle collega's aan te bevelen (*Hoofdstuk 6*).

Het onderzoeksnetwerk dat we geanalyseerd hebben om aanbevelingen te genereren werd afgeleid uit de DSpace database die publicaties en hun metadata van een lokale instelling bevat. Voor elk artikel in de database hebben we de co-auteurs bepaald, die onderling een co-auteurrelatie in een co-auteurschapsnetwerk vormen, en hebben we de sleutelwoorden bepaald. Het systeem gebruikt betweenness centrality om machtige collega's in het co-auteurschapsnetwerk te identificeren. Om gelijkgestemde mede-auteurs te bepalen gebruikt het systeem de gelijkenis tussen de sleutelwoorden die de auteurs gebruiken om hun documenten te beschrijven.

Het aanbevelingsalgoritme neemt het gewogen gemiddelde van zowel de betweenness van en de trefwoordgelijkenis van co-auteurs tot een bepaalde gebruiker. Aan gebruikers werden twee lijsten voorgelegd: één met alleen maar nieuwe co-auteurs, en één met nieuwe én bestaande co-auteurs. De resultaten toonden aan dat gebruikers de voorkeur geven aan een aanbeveling die bestaat uit bestaande co-auteurs, omdat ze relatief onbekend zijn met het werk van de co-auteurs waar ze niet nog mee samengewerkt hebben.

Het COCOON systeem werd opgevolgd door COCOON CORE. En net als COCOON, biedt het een middel om een praktische kijk op samenwerking (*Hoofdstuk 7*) te nemen. COCOON CORE richt zich op de verdere zelfbeschikking van de gebruikers door hen de gelegenheid te geven om de balans tussen het vinden van machtige collega's en gelijkgestemde collega's zelf aan te passen. Ook presenteren we auteurpagina's om meer inzicht te geven in wat auteurs schrijven, wat hun outputkwaliteit is, en hoe auteurs aan elkaar gerelateerd zijn. Tot slot presenteren we de sleutelwoordpagina's die de kwaliteit van sleutelwoorden aangeven en hoe ze gerelateerd zijn aan elkaar.

We hebben een evaluatie uitgevoerd onder onderzoekers van de instelling die de DSpace database herbergt en de resultaten toonden aan dat de deelnemers de mogelijkheid om zelf het aanbevelingsalgoritme aan te passen, goed waarderen. In het algemeen scoorden de aanbevelingen matig positief. COCOON CORE werd ook getest op de gebruiksvriendelijkheid. Het scoorde matig positief op bruikbaarheid, en met name de leerbaarheidscores waren hoog. Toekomstige werkzaamheden voor COCOON CORE moeten zich richten op de integratie van haar diensten.

Conclusie

Dit proefschrift is gericht op interpersoonlijke en intrapersoonlijke problemen in coöperatienetwerken. We hebben ons specifiek gericht op het oplossen van begrensde rationaliteit door middel van het aansturen van het beslissingsproces van de individu die op zoek is naar nieuwe samenwerking in zijn of haar netwerk. Ook hebben we ons gericht op het verminderen van de informatieoverdaad die mensen normaal gesproken tegenkomen als ze in hun netwerk zoeken naar waardevolle collega's. Elk hoofdstuk in dit proefschrift is gericht op het oplossen van een specifiek deelprobleem dat men tegenkomt in het stap-voor-stap proces dat nodig is om met succes een systeem in te voeren dat de samenwerking in netwerken ondersteunt.

De resultaten tonen aan dat een systeem dat machtige en gelijkgestemde collega's voor samenwerking aanbeveelt wordt gewaardeerd onder de gebruikers en dus potentieel heeft. In dit tijdperk van sociale media kan het bijzonder interessant zijn om verder onderzoek te doen in de richting van netwerkgebaseerde aanbevelingssystemen. Vanuit het perspectief van onderzoek naar sociale netwerken is het tijd om de volgende stap te maken naar een theorie over sociale netwerken die interventies informeert in plaats van louter de analyse van sociale netwerken.

Acknowledgements

Acknowledgements

I am grateful to my supervisor Peter Sloep for giving me the opportunity to learn. I can imagine that when I applied for this position, you may have had a look at my marks, which were definitely not stellar. I also remember that you were told that I was not that much of a sociable person. In fact I am, but I just need some time to adjust to my surroundings. Luckily you could see through all that. The very fact that you could, characterises what I most value in you. You are kind-hearted, easily approachable, and you are able to identify people's strengths and put them to use. You are one of few persons I feel I can rely on.

I am also thankful to my daily supervisor Marlies Bitter-Rijpkema. I remember the numerous times that we met to discuss my progress. One of the questions you kept repeating was "How does this relate to your project?," This reflects how you always kept an overview of where I was heading, in spite of my various (and countless) research detours. We both have a very different way of working. I tend to finish things quickly, but not so thoroughly, whereas you have a more scientific approach of thoroughly examining, in this case, our work. They are complementary, and I surely profited from your approach. It can also be frustrating, and I think you sometimes may have felt that you could not get through to me, although you never complained. Thank you.

I would like to thank Sibren Fetter for guiding me within CELSTEC during the first year of my PhD. Thank you for listening to me and critically assessing my thoughts. Thank you and Nicole for letting me stay over at your place every now and then, when I had early appointments at work, which I couldn't make in time if I had to travel from Amsterdam. Thank you for introducing me to Halo and Battlefield (Xbox videogames) and, not entirely to my taste, the Mango song (Weebl's Stuff, 2008). Thank you for eating my chocolate. Besides being a good colleague, you proved to be a good friend.

Special thanks to Mieke Haemers, the secretary who is actually much more than a secretary. You helped me find my way through the Open University, CELSTEC, Limburg and my PhD trajectory. You supported me in numerous ways, such as booking flights, train tickets and hotels, quelling my chaos and proofreading this thesis. I value your straight way of saying things.

I would like to thank my colleagues from CELSTEC and especially the Learning Networks team. Kamakshi Rajagopal, Adriana Berlanga, Francis Brouns, Ellen Rusman, Hendrik Drachsler, Jan van Bruggen, Peter van Rosmalen, Slavi Stoyanov, Kees Pannekeet, Marjo Rutjens, Amy Hsiao, and all the others I have forgotten to mention: thank you for your support and understanding, eating my pastry, visiting the meetings, discussions, ideas, etc.

Last but definitely not least, I would like to thank my wife, Elisabeth Uijttenbroek. You played a major role in the development of this thesis, both from a content perspective and a personal perspective. I think you deserve a major part in my

acknowledgement, but I could have written a book about how thankful I am for having you. I will condense the story a bit, but I feel I need to write this, regardless of what others may think about whether such a ‘long’ story belongs here or not. You have had and will have a great influence on my life.

You gave new direction to my life when we met about eight years ago. You already finished your law study, and you were about to finish your second law study. I was that ‘punk’ from Duivendrecht (a village on the outskirts of Amsterdam) that woke up late, watched every soccer game and sitcom on the telly, had few and strange clothes, interestingly had little to no money (although I did work), and woke up late. I told you I was in my third year of studying Artificial Intelligence. ‘Theoretically’ I was. Actually, I couldn’t make up my mind about what to study, and had been switching back and forth between Information Sciences, Psychology and Artificial Intelligence for the past two years. One could imagine how it would have sounded that I said I was in the third year of my study.

So we made a plan, as would we various times later on in our life. I would have two years of full focus on my study, without having to work. And you would work full-time, provided that I would switch back to studying Artificial Intelligence and gain a serious number of ECTS points. The plan proved to be successful, in contrast to many of our subsequent plans. I finished my study three years later and in the years after my study, I became increasingly interested in Artificial Intelligence and especially its applications.

One of our plans was to buy a new house at the start of my PhD, to live closer to the Open University in Heerlen, which is quite a distance from Almere where we lived at that time. The house and its seller proved to be not as reliable as we expected, resulting in losing a considerable sum of money and gaining a lot of stress. This was not an eligible reason, but it did give people at the Free University of Amsterdam the opportunity to get rid of you. Despite having more publications in one-and-a-half year than I have after four years. I feel like I am standing and working where you belong. You are the woman with a plan (it does rhyme as long as you do not pronounce it). You like reading and writing more than I do, and I think you are smarter than I am. Without a doubt, you are more of a researcher than I am, although you never had the opportunity to really show it. If I could ever exchange my PhD for you to have a second chance, I would do so instantly.

Curriculum Vitae

Curriculum Vitae

Rory Sie, MSc. (1982) is a PhD candidate at the Open Universiteit in the Netherlands. After graduating in artificial intelligence in 2007, he started working at the Free University of Amsterdam, where he worked at the Knowledge Representation and Reasoning group of the Computer Science department. In 2008, he joined the Centre for Learning Sciences and Technologies (CELSTEC) at the Open Universiteit in the Netherlands as a PhD candidate. From September 2012 onward, he will be working at the Wetenschappelijk Centrum Leraren Onderzoek (LOOK) at the Open Universiteit in the Netherlands.

At CELSTEC, Open Universiteit, he was involved in the EU FP7 project *IdSpace* about tooling for creativity and innovation. Within the IdSpace project, he worked on formalising and combining creative techniques and pedagogical strategies to support online, distant collaboration between new product designers. Also, he was involved in the evaluation of the IdSpace online collaboration environment. He is former chair and founder of the PhD council of the Open Universiteit in the Netherlands and former chair of the PhD council of the Dutch Research School for Information Knowledge Systems (SIKS).

His current focus is on cooperation networks (e.g. learning, innovation and research networks), and how we can use social network analysis and game theoretic solution concepts to foster successful cooperation. Particularly, he has been working on the COCOON project to support various stages of cooperation in networks, Firstly, he worked on the *COCOON Personal Learning Network Identification tool*, in which learners can identify their personal learning network from an ego perspective. Secondly, he worked on the *COCOON CORE* system that analyses a cooperation network and recommends valuable peers. Finally, he developed the *Tweetstorm* by which knowledge can be quickly elicited and analysed by employing the microblogging website Twitter to brainstorm dispersedly. In a more general sense, he is interested in how we can apply techniques from artificial intelligence (multi- agent systems, intelligent virtual agents, semantic web) to education, learning and cooperation. Other interests include the science of science and bread baking.

SIKS Dissertations Series

SIKS Dissertation Series

The complete list of SIKS dissertations (from 1998 onward) is available at http://www.siks.nl/dissertations.php .

====
2009
====

2009-01
Rasa Jurgelenaite (RUN)
Symmetric Causal Independence Models

2009-02
Willem Robert van Hage (VU)
Evaluating Ontology-Alignment Techniques

2009-03
Hans Stol (UvT)
A Framework for Evidence-based Policy Making Using IT

2009-04
Josephine Nabukenya (RUN)
Improving the Quality of Organisational Policy Making using Collaboration Engineering

2009-05
Sietse Overbeek (RUN)
Bridging Supply and Demand for Knowledge Intensive Tasks - Based on Knowledge, Cognition, and Quality

2009-06
Muhammad Subianto (UU)
Understanding Classification

2009-07
Ronald Poppe (UT)
Discriminative Vision-Based Recovery and Recognition of Human Motion

2009-08
Volker Nannen (VU)
Evolutionary Agent-Based Policy Analysis in Dynamic Environments

2009-09
Benjamin Kanagwa (RUN)
Design, Discovery and Construction of Service-oriented Systems

2009-10
Jan Wielemaker (UVA)
Logic programming for knowledge-intensive interactive applications

2009-11
Alexander Boer (UVA)
Legal Theory, Sources of Law & the Semantic Web

2009-12
Peter Massuthe (TUE, Humboldt-Universitaet zu Berlin)
Operating Guidelines for Services

2009-13
Steven de Jong (UM)
Fairness in Multi-Agent Systems

2009-14
Maksym Korotkiy (VU)
From ontology-enabled services to service-enabled ontologies (making ontologies work in e-science with ONTO-SOA)

2009-15
Rinke Hoekstra (UVA)
Ontology Representation - Design Patterns and Ontologies that Make Sense

2009-16
Fritz Reul (UvT)
New Architectures in Computer Chess

2009-17
Laurens van der Maaten (UvT)
Feature Extraction from Visual Data

2009-18
Fabian Groffen (CWI)
Armada, An Evolving Database System

2009-19
Valentin Robu (CWI)
Modeling Preferences, Strategic Reasoning and Collaboration in Agent-Mediated Electronic Markets

2009-20
Bob van der Vecht (UU)

Adjustable Autonomy: Controling Influences on Decision Making

2009-21
Stijn Vanderlooy (UM)
Ranking and Reliable Classification

2009-22
Pavel Serdyukov (UT)
Search For Expertise: Going beyond direct evidence

2009-23
Peter Hofgesang (VU)
Modelling Web Usage in a Changing Environment

2009-24
Annerieke Heuvelink (VUA)
Cognitive Models for Training Simulations

2009-25
Alex van Ballegooij (CWI)
"RAM: Array Database Management through Relational Mapping"

2009-26
Fernando Koch (UU)
An Agent-Based Model for the Development of Intelligent Mobile Services

2009-27
Christian Glahn (OU)
Contextual Support of social Engagement and Reflection on the Web

2009-28
Sander Evers (UT)
Sensor Data Management with Probabilistic Models

2009-29
Stanislav Pokraev (UT)
Model-Driven Semantic Integration of Service-Oriented Applications

2009-30
Marcin Zukowski (CWI)
Balancing vectorized query execution with bandwidth-optimized storage

2009-31
Sofiya Katrenko (UVA)
A Closer Look at Learning Relations from Text

2009-32
Rik Farenhorst (VU) and Remco de Boer (VU)
Architectural Knowledge Management: Supporting Architects and Auditors

2009-33
Khiet Truong (UT)
How Does Real Affect Affect Affect Recognition In Speech?

2009-34
Inge van de Weerd (UU)
Advancing in Software Product Management: An Incremental Method Engineering Approach

2009-35
Wouter Koelewijn (UL)
Privacy en Politiegegevens; Over geautomatiseerde normatieve informatie-uitwisseling

2009-36
Marco Kalz (OUN)
Placement Support for Learners in Learning Networks

2009-37
Hendrik Drachsler (OUN)
Navigation Support for Learners in Informal Learning Networks

2009-38
Riina Vuorikari (OU)
Tags and self-organisation: a metadata ecology for learning resources in a multilingual context

2009-39
Christian Stahl (TUE, Humboldt-Universitaet zu Berlin)
Service Substitution -- A Behavioral Approach Based on Petri Nets

2009-40
Stephan Raaijmakers (UvT)
Multinomial Language Learning: Investigations into the Geometry of Language

2009-41
Igor Berezhnyy (UvT)
Digital Analysis of Paintings

2009-42
Toine Bogers (UvT)
Recommender Systems for Social Bookmarking

2009-43
Virginia Nunes Leal Franqueira (UT)
Finding Multi-step Attacks in Computer Networks using Heuristic Search and Mobile Ambients

2009-44
Roberto Santana Tapia (UT)
Assessing Business-IT Alignment in Networked Organizations

2009-45
Jilles Vreeken (UU)
Making Pattern Mining Useful

2009-46
Loredana Afanasiev (UvA)
Querying XML: Benchmarks and Recursion

====
2010
====

2010-01
Matthijs van Leeuwen (UU)
Patterns that Matter

2010-02
Ingo Wassink (UT)
Work flows in Life Science

2010-03
Joost Geurts (CWI)
A Document Engineering Model and Processing Framework for Multimedia documents

2010-04
Olga Kulyk (UT)
Do You Know What I Know? Situational Awareness of Co-located Teams in Multidisplay Environments

2010-05
Claudia Hauff (UT)

Predicting the Effectiveness of Queries and Retrieval Systems

2010-06
Sander Bakkes (UvT)
Rapid Adaptation of Video Game AI

2010-07
Wim Fikkert (UT)
Gesture interaction at a Distance

2010-08
Krzysztof Siewicz (UL)
Towards an Improved Regulatory Framework of Free Software. Protecting user freedoms in a world of software communities and eGovernments

2010-09
Hugo Kielman (UL)
A Politiele gegevensverwerking en Privacy, Naar een effectieve waarborging

2010-10
Rebecca Ong (UL)
Mobile Communication and Protection of Children

2010-11
Adriaan Ter Mors (TUD)
The world according to MARP: Multi-Agent Route Planning

2010-12
Susan van den Braak (UU)
Sensemaking software for crime analysis

2010-13
Gianluigi Folino (RUN)
High Performance Data Mining using Bio-inspired techniques

2010-14
Sander van Splunter (VU)
Automated Web Service Reconfiguration

2010-15
Lianne Bodenstaff (UT)
Managing Dependency Relations in Inter-Organizational Models

2010-16

Sicco Verwer (TUD)
Efficient Identification of Timed Automata, theory and practice

2010-17
Spyros Kotoulas (VU)
Scalable Discovery of Networked Resources: Algorithms, Infrastructure, Applications

2010-18
Charlotte Gerritsen (VU)
Caught in the Act: Investigating Crime by Agent-Based Simulation

2010-19
Henriette Cramer (UvA)
People's Responses to Autonomous and Adaptive Systems

2010-20
Ivo Swartjes (UT)
Whose Story Is It Anyway? How Improv Informs Agency and Authorship of Emergent Narrative

2010-21
Harold van Heerde (UT)
Privacy-aware data management by means of data degradation

2010-22
Michiel Hildebrand (CWI)
End-user Support for Access to\\ Heterogeneous Linked Data

2010-23
Bas Steunebrink (UU)
The Logical Structure of Emotions

2010-24
Dmytro Tykhonov
Designing Generic and Efficient Negotiation Strategies

2010-25
Zulfiqar Ali Memon (VU)
Modelling Human-Awareness for Ambient Agents: A Human Mindreading Perspective

2010-26
Ying Zhang (CWI)

XRPC: Efficient Distributed Query Processing on Heterogeneous XQuery Engines

2010-27
Marten Voulon (UL)
Automatisch contracteren

2010-28
Arne Koopman (UU)
Characteristic Relational Patterns

2010-29
Stratos Idreos(CWI)
Database Cracking: Towards Auto-tuning Database Kernels

2010-30
Marieke van Erp (UvT)
Accessing Natural History - Discoveries in data cleaning, structuring, and retrieval

2010-31
Victor de Boer (UVA)
Ontology Enrichment from Heterogeneous Sources on the Web

2010-32
Marcel Hiel (UvT)
An Adaptive Service Oriented Architecture: Automatically solving Interoperability Problems

2010-33
Robin Aly (UT)
Modeling Representation Uncertainty in Concept-Based Multimedia Retrieval

2010-34
Teduh Dirgahayu (UT)
Interaction Design in Service Compositions

2010-35
Dolf Trieschnigg (UT)
Proof of Concept: Concept-based Biomedical Information Retrieval

2010-36
Jose Janssen (OU)

Paving the Way for Lifelong Learning; Facilitating competence development through a learning path specification

2010-37
Niels Lohmann (TUE)
Correctness of services and their composition

2010-38
Dirk Fahland (TUE)
From Scenarios to components

2010-39
Ghazanfar Farooq Siddiqui (VU)
Integrative modeling of emotions in virtual agents

2010-40
Mark van Assem (VU)
Converting and Integrating Vocabularies for the Semantic Web

2010-41
Guillaume Chaslot (UM)
Monte-Carlo Tree Search

2010-42
Sybren de Kinderen (VU)
Needs-driven service bundling in a multi-supplier setting - the computational e3-service approach

2010-43
Peter van Kranenburg (UU)
A Computational Approach to Content-Based Retrieval of Folk Song Melodies

2010-44
Pieter Bellekens (TUE)
An Approach towards Context-sensitive and User-adapted Access to Heterogeneous Data Sources, Illustrated in the Television Domain

2010-45
Vasilios Andrikopoulos (UvT)
A theory and model for the evolution of software services

2010-46
Vincent Pijpers (VU)
e3alignment: Exploring Inter-Organizational Business-ICT Alignment

2010-47
Chen Li (UT)
Mining Process Model Variants: Challenges, Techniques, Examples

2010-48
Withdrawn

2010-49
Jahn-Takeshi Saito (UM)
Solving difficult game positions

2010-50
Bouke Huurnink (UVA)
Search in Audiovisual Broadcast Archives

2010-51
Alia Khairia Amin (CWI)
Understanding and supporting information seeking tasks in multiple sources

2010-52
Peter-Paul van Maanen (VU)
Adaptive Support for Human-Computer Teams: Exploring the Use of Cognitive Models of Trust and Attention

2010-53
Edgar Meij (UVA)
Combining Concepts and Language Models for Information Access

====
2011
====

2011-01
Botond Cseke (RUN)
Variational Algorithms for Bayesian Inference in Latent Gaussian Models

2011-02
Nick Tinnemeier(UU)

Organizing Agent Organizations. Syntax and Operational Semantics of an Organization-Oriented Programming Language

2011-03
Jan Martijn van der Werf (TUE)
Compositional Design and Verification of Component-Based Information Systems

2011-04
Hado van Hasselt (UU)
Insights in Reinforcement Learning; Formal analysis and empirical evaluation of temporal-difference
learning algorithms

2011-05
Base van der Raadt (VU)
Enterprise Architecture Coming of Age - Increasing the Performance of an Emerging Discipline.

2011-06
Yiwen Wang (TUE)
Semantically-Enhanced Recommendations in Cultural Heritage

2011-07
Yujia Cao (UT)
Multimodal Information Presentation for High Load Human Computer Interaction

2011-08
Nieske Vergunst (UU)
BDI-based Generation of Robust Task-Oriented Dialogues

2011-09
Tim de Jong (OU)
Contextualised Mobile Media for Learning

2011-10
Bart Bogaert (UvT)
Cloud Content Contention

2011-11
Dhaval Vyas (UT)
Designing for Awareness: An Experience-focused HCI Perspective

2011-12
Carmen Bratosin (TUE)
Grid Architecture for Distributed Process Mining

2011-13
Xiaoyu Mao (UvT)
Airport under Control. Multiagent Scheduling for Airport Ground Handling

2011-14
Milan Lovric (EUR)
Behavioral Finance and Agent-Based Artificial Markets

2011-15
Marijn Koolen (UvA)
The Meaning of Structure: the Value of Link Evidence for Information Retrieval

2011-16
Maarten Schadd (UM)
Selective Search in Games of Different Complexity

2011-17
Jiyin He (UVA)
Exploring Topic Structure: Coherence, Diversity and Relatedness

2011-18
Mark Ponsen (UM)
Strategic Decision-Making in complex games

2011-19
Ellen Rusman (OU)
The Mind ' s Eye on Personal Profiles

2011-20
Qing Gu (VU)
Guiding service-oriented software engineering - A view-based approach

2011-21
Linda Terlouw (TUD)
Modularization and Specification of Service-Oriented Systems

2011-22
Junte Zhang (UVA)
System Evaluation of Archival Description and Access

2011-23
Wouter Weerkamp (UVA)
Finding People and their Utterances in Social Media

2011-24
Herwin van Welbergen (UT)
Behavior Generation for Interpersonal Coordination with Virtual Humans On Specifying, Scheduling and Realizing Multimodal Virtual Human Behavior

2011-25
Syed Waqar ul Qounain Jaffry (VU))
Analysis and Validation of Models for Trust Dynamics

2011-26
Matthijs Aart Pontier (VU)
Virtual Agents for Human Communication - Emotion Regulation and Involvement-Distance Trade-Offs in Embodied Conversational Agents and Robots

2011-27
Aniel Bhulai (VU)
Dynamic website optimization through autonomous management of design patterns

2011-28
Rianne Kaptein(UVA)
Effective Focused Retrieval by Exploiting Query Context and Document Structure

2011-29
Faisal Kamiran (TUE)
Discrimination-aware Classification

2011-30
Egon van den Broek (UT)
Affective Signal Processing (ASP): Unraveling the mystery of emotions

2011-31
Ludo Waltman (EUR)
Computational and Game-Theoretic Approaches for Modeling Bounded Rationality

2011-32
Nees-Jan van Eck (EUR)
Methodological Advances in Bibliometric Mapping of Science

2011-33
Tom van der Weide (UU)
Arguing to Motivate Decisions

2011-34
Paolo Turrini (UU)
Strategic Reasoning in Interdependence: Logical and Game-theoretical Investigations

2011-35
Maaike Harbers (UU)
Explaining Agent Behavior in Virtual Training

2011-36
Erik van der Spek (UU)
Experiments in serious game design: a cognitive approach

2011-37
Adriana Burlutiu (RUN)
Machine Learning for Pairwise Data, Applications for Preference Learning and Supervised Network Inference

2011-38
Nyree Lemmens (UM)
Bee-inspired Distributed Optimization

2011-39
Joost Westra (UU)
Organizing Adaptation using Agents in Serious Games

2011-40
Viktor Clerc (VU)
Architectural Knowledge Management in Global Software Development

2011-41
Luan Ibraimi (UT)
Cryptographically Enforced Distributed Data Access Control

2011-42
Michal Sindlar (UU)
Explaining Behavior through Mental State Attribution

2011-43
Henk van der Schuur (UU)
Process Improvement through Software Operation Knowledge

2011-44
Boris Reuderink (UT)
Robust Brain-Computer Interfaces

2011-45
Herman Stehouwer (UvT)
Statistical Language Models for Alternative Sequence Selection

2011-46
Beibei Hu (TUD)
Towards Contextualized Information Delivery: A Rule-based Architecture for the Domain of Mobile Police Work

2011-47
Azizi Bin Ab Aziz(VU)
Exploring Computational Models for Intelligent Support of Persons with Depression

2011-48
Mark Ter Maat (UT)
Response Selection and Turn-taking for a Sensitive Artificial Listening Agent

2011-49
Andreea Niculescu (UT)
Conversational interfaces for task-oriented spoken dialogues: design aspects influencing interaction quality

====
2012
====

2012-01
Terry Kakeeto (UvT)
Relationship Marketing for SMEs in Uganda

2012-02
Muhammad Umair(VU)
Adaptivity, emotion, and Rationality in Human and Ambient Agent Models

2012-03
Adam Vanya (VU)
Supporting Architecture Evolution by Mining Software Repositories

2012-04
Jurriaan Souer (UU)
Development of Content Management System-based Web Applications

2012-05
Marijn Plomp (UU)
Maturing Interorganisational Information Systems

2012-06
Wolfgang Reinhardt (OU)
Awareness Support for Knowledge Workers in Research Networks

2012-07
Rianne van Lambalgen (VU)
When the Going Gets Tough: Exploring Agent-based Models of Human Performance under Demanding Conditions

2012-08
Gerben de Vries (UVA)
Kernel Methods for Vessel Trajectories

2012-09
Ricardo Neisse (UT)
Trust and Privacy Management Support for Context-Aware Service Platforms

2012-10
David Smits (TUE)
Towards a Generic Distributed Adaptive Hypermedia Environment

2012-11
J.C.B. Rantham Prabhakara (TUE)
Process Mining in the Large: Preprocessing, Discovery, and Diagnostics

2012-12
Kees van der Sluijs (TUE)
Model Driven Design and Data Integration in Semantic Web Information Systems

2012-13
Suleman Shahid (UvT)
Fun and Face: Exploring non-verbal expressions of emotion during playful interactions

2012-14
Evgeny Knutov(TUE)
Generic Adaptation Framework for Unifying Adaptive Web-based Systems

2012-15
Natalie van der Wal (VU)
Social Agents. Agent-Based Modelling of Integrated Internal and Social Dynamics of Cognitive and Affective Processes.

2012-16
Fiemke Both (VU)
Helping people by understanding them - Ambient Agents supporting task execution and depression treatment

2012-17
Amal Elgammal (UvT)
Towards a Comprehensive Framework for Business Process Compliance

2012-18
Eltjo Poort (VU)
Improving Solution Architecting Practices

2012-19
Helen Schonenberg (TUE)
What's Next? Operational Support for Business Process Execution

2012-20
Ali Bahramisharif (RUN)
Covert Visual Spatial Attention, a Robust Paradigm for Brain-Computer Interfacing

2012-21
Roberto Cornacchia (TUD)
Querying Sparse Matrices for Information Retrieval

2012-22
Thijs Vis (UvT)
Intelligence, politie en veiligheidsdienst: verenigbare grootheden?

2012-23
Christian Muehl (UT)
Toward Affective Brain-Computer Interfaces: Exploring the Neurophysiology of Affect during Human Media Interaction

2012-24

Laurens van der Werff (UT)
Evaluation of Noisy Transcripts for Spoken Document Retrieval

2012-25
Silja Eckartz (UT)
Managing the Business Case Development in Inter-Organizational IT Projects: A Methodology and its Application

2012-26
Emile de Maat (UVA)
Making Sense of Legal Text

2012-27
Hayrettin Gurkok (UT)
Mind the Sheep! User Experience Evaluation & Brain-Computer Interface Games

2012-28
Nancy Pascall (UvT)
Engendering Technology Empowering Women

2012-29
Almer Tigelaar (UT)
Peer-to-Peer Information Retrieval

2012-30
Alina Pommeranz (TUD)
Designing Human-Centered Systems for Reflective Decision Making

2012-31
Emily Bagarukayo (RUN)
A Learning by Construction Approach for Higher Order Cognitive Skills Improvement, Building Capacity and Infrastructure

2012-32
Wietske Visser (TUD)
Qualitative multi-criteria preference representation and reasoning