# Examining the Interplay Between Universal Behavioural Tendencies, Online Social Networks and Social Capital

DOCTORAL THESIS

**Jayant Venkatanathan**
DOCTORATE IN INFORMATICS ENGINEERING
SPECIALTY: HUMAN-COMPUTER INTERACTION

SUPERVISOR
Evangelos Karapanos

CO-SUPERVISOR
Vassilis Kostakos

## Abstract

Interaction with others is fundamental to well-being, as it serves to fulfil our basic needs. Thus humans have various behavioural tendencies, patterns of behaviour that serve as strategies to fulfil these needs. Given the increasingly crucial role of online social networks on our communication and interaction, it is important to study these factors in the online context. In this thesis we explore how universal behavioural tendencies, i.e. behavioural tendencies that have been observed across cultures, affect our online interaction and how these in turn affect social capital. Focusing on disclosure behaviour and social network structure as proxies for online interaction behaviour, this work consists of three main components developed over four studies. Firstly, we attempt to understand how the tendency to reciprocate affects individuals' willingness to disclose information about themselves. Secondly, we study the interplay between individuals' disclosure patterns and their positions in the network. Finally, we study how individuals, along with their differences in universal behavioural tendencies, accrue social capital from the structure of their immediate networks. Key findings include: (1) People tend to reciprocate the disclosure of personal information, both when the initial disclosure is directed towards them, and also when it is broadcast and directed to nobody in particular, (2) The centrality of individuals in a social network is related to how much information they disclose, and how much others disclose to them, and (3) Online social network structure is related to social capital, and network structure and empathy play an interconnected role in the creation of social capital. The empirical findings, discussions and methodologies presented in this work will be useful for HCI and social science researchers studying the fundamental aspects of humans' use of social technologies.

**Keywords:** Online Social Networks; Privacy; Social Capital; Social Networks Analysis; Human Computer Interaction; Web Science.

## Resumo

A interação com os outros é essencial para o bem-estar, visto servir para satisfazer as nossas necessidades básicas. Portanto, os seres humanos têm várias tendências comportamentais, padrões de comportamento que servem como estratégias para satisfazer essas necessidades. Dada a importância crescente das redes sociais online na nossa comunicação e interação, é importante estudar estes fatores no contexto online. Na presente tese exploramos como as tendências comportamentais universais, i.e., as tendências comportamentais observadas em diferentes culturas afetam a nossa interação online e como estas, por sua vez, afetam o capital social. Concentrando-se na divulgação comportamental e na estrutura da rede social como representantes do comportamento interativo online, este trabalho apresenta três componentes principais desenvolvidas em 4 estudos. Primeiro, tentamos compreender de que forma a tendência para a reciprocidade afeta a vontade dos indivíduos de divulgarem informações sobre eles mesmos. Segundo, estudamos a interação entre os padrões de divulgação dos indivíduos e as suas posições na rede. Finalmente, estudamos de que forma os indivíduos, juntamente com as suas diferenças nas tendências comportamentais universais, acumulam capital social a partir da estrutura das suas redes imediatas. As principais conclusões incluem: (1) As pessoas retribuem a divulgação de informação pessoal não só quando esta é dirigida ao próprio, mas de igual forma se publicada num espaço publico acessível a qualquer pessoa, (2) A centralidade dos indivíduos numa rede social está relacionada com a quantidade de informações que divulga e que os outros lhes divulgam, e (3) A estrutura da rede social online está relacionada com o capital social, e a estrutura da rede e empatia desempenham um papel próximo na criação do capital social. Os resultados empíricos, discussões e metodologias apresentados neste trabalho serão úteis para os investigadores de HCI e ciências sociais que estudam os aspetos fundamentais da utilização humana das tecnologias sociais.



**Palavras chave:** Redes Sociais Online; Privacidade; Capital Social; Análise de Redes Sociais; Interação Humano-Computador; Ciência da Web.

## Acknowledgements

Much thanks go to Vassilis Kostakos and Evangelos Karapanos for the support and guidance they have offered me over the years. I am also grateful to colleagues and friends at the University of Madeira and the University of Oulu for their support and friendship.

This work was supported by the Portuguese Foundation for Science and Technology (FCT) grant CMU-PT/SE/ 0028/2008 (Web Security and Privacy).

## Statement of Attribution

This doctoral dissertation is based on material from the following journal (a), conference (b, c) and workshop (d) publications:

(a) Venkatanathan, J., Kostakos, V., Karapanos, E., Goncalves, J. (2014). Online disclosure of personally identifiable information with strangers: Effects of public and private sharing. Interacting with Computers, 26(6), 614-626. doi:10.1093/iwc/iwt058.

(b) Kostakos, V., Venkatanathan, J., Reynolds, B., Sadeh, N., Toch, E., Shaikh, S. A., Jones, S. (2011). Who's your best friend? Targeted privacy attacks in location-sharing social networks. In proceedings of Ubicomp 2011, Beijing, China, p. 177-186. doi: 10.1145/2030112.2030138.

(c) Venkatanathan, J., Karapanos, E., Kostakos, V., Goncalves, J. (2012). Network, Personality and Social Capital. In proceedings of ACM Web Science, Evanston, USA, p. 326-329. doi: 10.1145/2380718.2380760.

(d) Venkatanathan, J., Karapanos, E., Kostakos, V., Goncalves, J. (2013). A Network Science Approach to Modelling and Predicting Empathy. In adjunct proceedings ASONAM 2013, International Workshop on Web Behavior Analytics. Niagara Falls, Canada. doi: 10.1145/2492517.2500295.



I was the lead investigator for the work in (a). The study was inspired from prior discussions with Vassilis Kostakos and implemented by me. Vassilis Kostakos also contributed to the planning of the study. The writing of the manuscript was primarily undertaken by me, with contributions from Vassilis Kostakos. Evangelos Karapanos and Jorge Goncalves contributed to the refinement around various discussions in the manuscript.

The work in (b) was led by Vassilis Kostakos, to which I made significant contributions. This work was conceived by Vassilis Kostakos in discussions with Siraj Sheikh and

myself. These initial ideas were then developed jointly by Vassilis Kostakos, Siraj Sheikh, Bernardo Reynolds, Simon Jones and I, and tested on data prepared and made available by Eran Toch and Norman Sadeh. Vassilis Kostakos and I carried out the coding of the data into graph structures and the subsequent analyses. Vassilis Kostakos lead the writing of the manuscript, to which I contributed. All authors contributed to the editing and refinement of the manuscript.

I was the lead investigator for the work in (c) and (d). These studies were conceived and developed jointly by Evangelos Karapanos, Vassilis Kostakos and me. Jorge Goncalves assisted in the development of the software for data collection. I was responsible for the writing of the manuscript, to which Evangelos Karapanos and Vassilis Kostakos made significant contributions. All authors contributed to the editing and refinement of the manuscripts.

# Table Of Contents

# 1. Introduction

Interaction with those around us is fundamental to our well-being, as they constitute and serve to fulfil our basic needs as human beings. Thus over the course of individuals' lifetime, and perhaps over the ages as a collective, humans appear to acquire and develop various behavioural tendencies that serve as strategies to fulfil these needs. Various factors such as genetics, environment, social norms, and the very unique circumstances that each individual encounters over the course of their lifetime condition patterns of behaviour in us that are triggered in different situations. Thus, individuals carry these learned patterns of behaviour or behavioural tendencies into the social situations they encounter, and this affects how they respond to a particular situation.

Given that interaction with others is fundamental to our well-being and to the fulfilment of needs, it is inevitable that how an individual responds to different social situations will affect her well-being. For example, some individuals might be comfortable approaching a distant relative for help regarding a job, while others might not be comfortable doing so. Thus, the behavioural tendencies we carry with us into our interactions affect our behaviour in different situations, and this in turn affects our overall well-being.

#### *Universal Behavioural Tendencies*

For any given individual, behavioural tendencies can be numerous, and their nature can be very unique to that individual. However, across individuals we can abstract out and categorise tendencies that can be seen as common, that individuals share to different extents. For example, the tendency to reciprocate favours is a behavioural tendency that every known human society has been reported to subscribe to (Gouldner, 1960). Similarly, similar personality trait structures characterise individuals across different cultures (McCrae and Costa, 1997). We refer to such behavioural tendencies, that have been observed across cultures, as *universal behavioural tendencies*. The term "universal" is not meant to suggest that these behavioural tendencies are present equally, or even present to a significant extent, in all individuals, but rather that these tendencies are pervasive and can be found across cultures. Universal behavioural tendencies are abstractions that might not capture the unique way in which each individual is wired to behave. Their value, however, lies in the fact that they allow us to

study behavioural tendencies at a population level, and examine how these tendencies correlate with other factors across the population.

### *Online Interactions*

As we carry over our face to face interactions into the online, the behavioural characteristics that have governed our offline interactions over the ages now also influence the way in which we interact and are affected by online interactions. Firstly, behavioural tendencies affect the ways in which we interact and behave in online social networks. This can be seen from the fact that many of the universal behavioural tendencies observed in offline interactions are also found to play out in online interactions. For example, while the tendency to reciprocate a favour or a disclosure of personal information has been well documented in face to face interactions (See for eg. Gouldner, 1962), it has also been found that reciprocity takes place in online forums when it comes to the disclosure of personal information. As another example, personality traits such as extraversion and conscientiousness affect usage of various features and behaviour on Facebook (Bachrach et al, 2012). The basic idea of behavioural tendencies affecting interaction behaviour can also be turned around to predict behavioural tendencies such as personality traits and validate theories of social interaction from interaction data in online social networks. For example, interaction data from online social networks in which links between individuals can be abstracted as being either positive (friendly) or negative (antagonistic) have been used to validate theories of signed networks from social psychology in their respective online contexts (Leskovec et al, 2010). Thus, behavioural tendencies are related to online social network behaviour in complex and varied ways.

Secondly, online social networks are having a profound effect on our communication behaviours and social well-being. While most social network sites support the maintenance of pre-existing social networks, others help strangers connect based on shared interests, political views, or activities (Boyd and Elison, 2007). These systems have an effect on our well-being (Elison et al, 2007), and have a number of social implications, leading researchers to argue that users feel more connected to their ties on a daily basis (Deters & Mehl, 2013). Thus our usage and interaction in online social networks affects overall well-being and the fulfilment of needs.

## 1.1 The Research Question

By synthesising the above discussed arguments, we can represent the joint effect of interactions in the offline and in online social networks on individuals and societies using a basic map. This map expresses the relationship between behavioural tendencies, the actual interaction and sharing behaviour in online social networks and

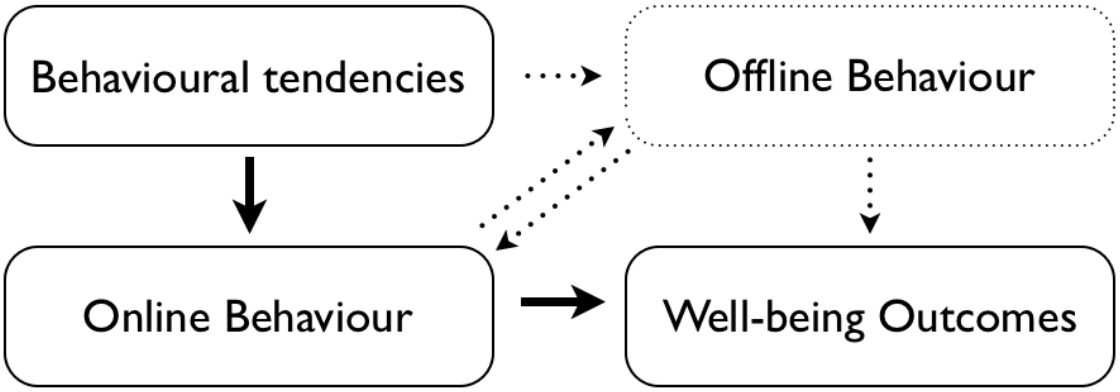


Figure 1 : Basic map representing the influence of online social networks on individuals and society

correspondingly in offline interactions, and the outcomes of well-being that result from these behaviours (Figure 1). The focus of this thesis is on online behaviour rather than offline behaviour. However, it is difficult to precisely separate online behaviour from offline behaviour, and the effect that these two have on well-being. In fact, one of the keys ways in which online social networks can be said to affect our well-being is by facilitating offline interactions. For example, online social networks allow us to co-ordinate offline events and meetings, or maintain latent offline ties. While there exist social networking sites and interactions in these sites that are purely between online ties, popular social networking sites such as Facebook are primarily concerned with people who already know each other, and use the Internet as one way of keeping their existing offline connections alive (Boyd and Ellison, 2007). It is for this reason that we also include offline behaviour in our conceptual map.

In this thesis we restrict our examination of behavioural tendencies to universal behavioural tendencies. Given the current large scale availability of social network data, universal behavioural tendencies have an advantage that they can easily be analysed as variables along with other social network metrics and metrics of well-being to understand phenomena at the level of the population. For well-being, in this thesis we

will restrict our focus to social capital. Social capital broadly refers to the resources and support that an individual has access to on account of his or her social ties. While one can consider other aspects of well-being such as health, income, happiness, etc., for this thesis we consider social capital. Correspondingly, we largely focus on interaction behaviours of individuals rather than all kinds of behaviours, as interaction behaviour is primarily related to social capital.

The purpose of this thesis, therefore, is to examine the interplay between universal behavioural tendencies, online interaction behaviour and social capital. Thus, we can state the overarching research question of the thesis as follows:

*Research Question - How do universal behavioural tendencies influence online interaction behaviour, and how do they both in turn affect social capital?*

Given the profound effects of online social networks on individuals and communities, it is crucial to examine these phenomena in detail. Any attempt to begin to satisfactorily address this question, given the richness and diversity of the use of online social networks, would require that we probe these issues from multiple perspectives and through combining multiple disciplines. Of course, given its very fundamental nature, this research question has countless facets to it that can be studied. The following section will provide a clearer picture of the aspects of the overarching research question that we deal with and the general standpoint from which we approach them, and the subsequent chapter (Background) will introduce the precise problems that we study.

## 1.2 Approach of the thesis

#### *Proxies for online interaction behaviour*

Online interaction behaviours include all kinds of online behaviours that individuals partake in, when interacting with others. There are numerous ways in which individuals can interact online ranging from exchanging private messages to indications, such as of approval or support, through “likes” or “upvotes”, and numerous ways in which interactions can be measured such as through the number of messages exchanged or the total time spent on the site or a particular feature of the site that facilitates interactions. In our work we narrow down our focus to two fundamental proxies that are indicative of

interactions between individuals: i) *disclosure* of personal information, and ii) network *structure*.

Disclosure of personal information, commonly referred to in the literature as self-disclosure, is an important indicator of the depth of interactions, as inherent in it is the element of trust. Self-disclosures are also an integral aspect of the process of the formation and strengthening of ties. It is for this reason that the effect of behavioural tendencies on the disclosure of personal information would be interesting and fruitful to explore. Further, self-disclosure is also related to social capital for the same reasons, as is discussed in Ellison et al. (2011).

Network structure refers to the patterns in which a user befriends others over time. Most studies in the HCI literature investigating the effects of online interactions conduct their analysis using two primary sources of data: data collected from users (such as questionnaire, attitude and behavioural data) and data collected from user interface mechanisms (such as usage logs and content analysis). The use of network structure as a third source of data has largely been overlooked, perhaps due to its rather implicit nature. It effectively acts as a backdrop against which social network activity takes place. Network structure can be a valuable source of data in the study of social capital particularly as a body of work in the social sciences has suggested a relationship between network structure and social capital (Eg. Granovetter, 1960; Burt, 1995; Rosenthal, 1996). It is for these reasons that we chose to focus on network structure as our second proxy for online interactions.

#### *Population-level perspective*

The approach we adopt in this thesis is a population level approach, whereby we measure effects over the population, rather than examine individuals in detail. The large-scale proliferation and adoption of online social networks provides a wealth of information of interactions between users in these online spaces. We seek to leverage these sources of data for understanding the interplay between behavioural tendencies, online interactions and social capital across the population. In addition to the availability of online interaction data, the abstraction of behavioural tendencies into universal behavioural tendencies facilitates this pursuit. The limitation of this approach however is that we can only study the effects of these variables in general over the population. We will not be able distinguish the finer details of how behavioural

tendencies manifest in a particular individual, or the mechanisms that might be affecting individuals who appear as outliers in our analyses. Such investigations are outside the scope of our current work.

***"Sense-Making" Driven Approach***

In this thesis, and particularly in the work presented in chapters 5 and 6 where we study questions around social capital and empathy in the context of online social networks, our investigation does not seek to solve any particular problem at hand, but rather to investigate an important problem space we have identified, by studying and mapping the occurring phenomena. Our driving motivation can be said to be akin to the goal of a "general sense-making of the world" (Bannon, 2011). As a result, our focus is not on the immediate practical applications or implications for the design of technological systems. Our primary motivation is to investigate fundamental phenomena, in the spirit of asking "What is really happening here?".

Of course, given that we are dealing with certain fundamental aspects of humans' use of technologies, our research is likely and is intended to impact our understanding and design of social technologies. Thus, where we are able to see with some clarity how this work can impact design, we have discussed those implications. However at the stage of the work developed thus far regarding the relationship between network structure, social capital and empathy in chapters 5 and 6 of the thesis, we feel it is premature to draw implications for the design of social technologies.

Overall, the work in this thesis lies at the intersection of the fields of Human-Computer Interaction (HCI), and the relatively nascent field of Web Science (Hendler et al, 2008), which studies the impact of the web on society and vice versa (see Hooper & Dix, 2012, for a discussion on the comparison of these two fields). Since our motivation is to work with and understand the fundamental aspects of humans' use of technologies so that this may subsequently affect the way we perceive and tackle problems, our approach is akin to "Pastuer's Quadrant" as described by Donald Stokes (1997): we seek to conduct use-inspired basic research which can contribute to fundamental scientific understanding while also provide practical value.

Having laid out the motivation and scope of the thesis, in the next chapter we provide a brief account of some related work around the conceptual map presented above (Section

1.2), and this will provide the background to lead us into the particular research problems examined in the subsequent chapters of the thesis.

**Chapter References**


Bachrach, Y., Kosinski, M., Graepel, T., Kohli, P., & Stillwell, D. (2012). Personality and patterns of Facebook usage. In Proceedings of the 3rd Annual ACM Web Science Conference (pp. 24-32). ACM.

Bannon, L. (2011). Reimagining HCI: toward a more human-centered perspective. interactions, 18(4), 50-57.

Boyd, D. M., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. Journal of Computer-Mediated Communication, 13(1), 210-230.

Burt, R. S. (1995). Structural Holes: The Social Structure of Competition. Harvard University Press.

Deters, F. G., & Mehl, M. R. (2013). Does posting Facebook status updates increase or decrease loneliness? An online social networking experiment. Social Psychological and Personality Science, 4, 579–586.

Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The benefits of Facebook "friends:" Social capital and college students' use of online social network sites. Journal of Computer-Mediated Communication, 12(4), 1143-1168.

Ellison, N., Vitak, J., Steinfield, C., Gray, R., and Lampe, C. (2011). Negotiating privacy concerns and social capital needs in a social media environment. Privacy online: Perspectives on privacy and self-disclosure in the social web (eds. S. Trepte & L. Reinecke). Springer, New York, 19-32.

Gouldner, A.W. (1960) The norm of reciprocity: a preliminary statement. Am. Soc. Rev., 25, 161–178.

Hendler, J., Shadbolt, N., Hall, W., Berners-Lee, T., & Weitzner, D. (2008). Web science: an interdisciplinary approach to understanding the web. Communications of the ACM, 51(7), 60-69.

Hooper, C. J., & Dix, A. (2012). Web Science and Human-Computer Interaction: When Disciplines Collide. In Proceedings of the 3rd Annual ACM Web Science Conference (pp. 128-136). ACM.

Leskovec, J., Huttenlocher, D., & Kleinberg, J. (2010). Predicting positive and negative links in online social networks. In Proceedings of the 19th international conference on World wide web (pp. 641-650). ACM.

McCrae, R. R., & Costa Jr, P. T. (1997). Personality trait structure as a human universal. American psychologist, 52(5), 509.

Rosenthal, E. A. (1996). Social Networks and Team Performance. Ph.D. Dissertation, Graduate School of Business, University of Chicago.

Stokes, D. E. (1997). Pasteur's Quadrant – Basic Science and Technological Innovation, Brookings Institution Press.

# 2. Background

We mapped out and presented our conceptual framework with the objective of framing online interactions in a broader perspective. By doing so, we aim to make explicit the more fundamental aspects of the human condition and how these are interlaced with online behaviour. As we present our work in this thesis, we attempt to hold this larger picture as we examine online behaviour. Nevertheless, the framework itself is fundamental and basic, and a body of prior work on online social networks has implicitly looked into different aspects of the framework. Here we briefly touch upon prior work that relates to and lends support to our conceptual framework.

## 2.1 Online Behaviour, and its relation to behavioural tendencies

There have been a number of recent studies that touch upon the association between universal behavioural tendencies, through concepts such as personality, and actual behaviour in online social networks. Differences between individuals in terms of personality traits and skills is related to how they use social networking systems (Anderson et al, 2012). For example, one study suggests that Facebook users who have higher levels of neuroticism tend to upload more photos of themselves, but upload fewer photos in general (Amichai-Hamburger and Vinitzky, 2010). In addition, demographic features such as gender also influence privacy attitudes (Stutzman & Kramer-Duffield, 2010). Thus overall we can expect that individual differences affect behaviour and usage in online social networks.

While focussing on the differences in behavioural tendencies between individuals can be one approach to understanding online behaviour, a complementary approach would be to look at commonalities across individuals - the common ways in which people tend to behave in different contexts. In relation to this, research looking at online social network behaviour can also draw from a wealth of research in traditional social science and social psychology. For example, Leskovek et al (2010) draw from social science theories such as Heider's theory of structural balance (Heider, 1958) to examine the interplay between positive (friendly) and negative (antagonistic) links in online social media. Similarly, Gilbert and Karahalios (2009) have attempted to draw from early work in social science such as Granovetter's (1973) work on the strength of ties to predict tie strength from interactions between individuals on Facebook.

An important aspect of online interaction is the disclosure of personal information, as this builds trust and brings individuals closer (Rubin, 1975) but also makes individuals vulnerable and has implications in terms of privacy. In terms of disclosure of information, we can draw from research in the context of face to face relations on the relationship between fundamental behavioural tendencies and self disclosure. For example, the fundamental behavioural tendency to reciprocate the disclosure of personal information in face to face information (see for eg. Archer and Berg, 1978; Collins and Miller, 1994) can provide insights into similar behaviour in online interactions. Thus, reciprocity of self disclosure has been reported in online forums where "self-disclosure" was measured by adding together instances of disclosures of facts, thoughts, and feelings about oneself (Barak and Gluck-Ofri, 2007). There has been recent work on understanding public disclosures. For example, on Facebook, disclosure shared privately is perceived to be more intimate than disclosure shared publicly (Bazarova, 2012a).

Prior work on social networks may also be used to derive hypotheses about, for example, who is likely to share information with whom on a social network. For instance, Petronio's theory of Communications Privacy Management (CPM) describes an iterative process of rule development, boundary coordination and boundary turbulence (Petronio, 2002). Rule development can be defined as the process of developing regulations about who to tell what. These regulations guide our everyday disclosures, and are a function of our context and disclosure goals. Ties of differing strength have varying disclosure norms, thus Stutzman and Kramer-Duffield (2010) theorize rule development is a function of network composition. For example, a network that is more heavily focused on strong ties may require higher levels of privacy, as disclosures among strong ties are more personal in nature (Wellman and Wortley, 1990). This suggests that network structure may be used as a basis for attempting to predict disclosures amongst individuals. Thus overall, there is ample support suggesting that behavioural tendencies of individuals are closely related to usage and interaction behaviour in online social networks.

## 2.2 Online Behaviour and Social Capital

Social Networking Sites (SNSs) are becoming increasingly prevalent and important instruments for users to manage their social life. As of September 2012, there were over one billion monthly active members on Facebook, more than half of whom access Facebook via a mobile device (The Wall Street Journal, 2012). Growing evidence suggests that SNSs have become important tools for managing relationships with a large and often heterogeneous network of people who provide social support (Boyd and Ellison, 2007; Ellison et al, 2007). Given this widespread and growing use of online social networks, a careful assessment of whether and how it affects users' well-being is crucial. In other words, how do users benefit from online social networks?

In order to operationalise the well-being outcomes and benefits that users' get from online social networks, we use the concept of social capital. The terms social capital has broadly been used to refer to the value of relationships between individuals and groups, and the resources that an individual has access to on account of his or her social ties (Portes, 1998; Putnam, 2000). There are 3 types of findings often reported in prior studies of Internet use and social capital (Steinfield et al, 2012).

(1) Internet use promotes social capital - The studies suggest that greater internet use is associated with the formation of meaningful relationships which in turn serve as a source of social support (Best and Kruger, 2006; Hampton and Wellman, 2002).
(2) Internet use can diminish social capital - The basic argument here is that distant online contacts cannot provide the same types of social support as physically proximate ones. These studies suggest that since time spent interacting with people online replaces time spent in face to face interactions, greater internet use diminishes social capital (Kraut et al, 1998; Nie, 2001).
(3) Internet use reinforces offline interactions and can therefore supplement social capital development - This takes a more nuanced view and sees the internet as a supplement rather than a substitute to other methods of communication (Quan Haase and Wellman, 2004).

Drawing support for the third perspective above, there have been substantial findings which suggest that SNSs blend the offline and online, rather than operating as independent and isolated systems of activity (Ellison et al, 2007; Lampe et al, 2008; Subrahmanyam et al, 2008). Thus, SNSs such as Facebook are typically used as a

means of building and maintaining relationships involving those with whom users share "some common offline element" (Boyd & Ellison, 2007). OSN's facilitate the interaction between spatially dispersed individuals and thus allows for the maintenance of ties that may have otherwise gone dormant (Ellison et al, 2007). Overall, there is evidence to support that the use of Facebook was found to be positively related to bonding social capital, which refers to the benefits obtained due to emotionally close relationships such as family and close friends, as well as bridging social capital, the benefits obtained from a looser network of weak ties (Ellison et. al., 2007). Research has also considered how individuals' differences in terms of personality traits or skills affect how they use and benefit from these systems (Anderson et al, 2012).

While earlier studies (c.f. Ellison et al, 2007) clumped together all activity on the social network and examine the relationship between overall time spent on the social network and social capital outcomes, more recent work began to distinguish between different activities on the social network for the effects they had on social capital (Eg. Burke et al, 2011; Yoder and Stutzman, 2011). Here we begin to see more careful examination of actual usage and interaction behaviour on the one hand and well-being outcomes on the other. For example, Burke et al (2011) separately examine the effect of passive content consumption and the effect of directed person to person exchanges. Directed person to person exchanges were found to be associated with increased bridging social capital, but only those with lower social communication skills experience higher social capital through content consumption (Burke et al, 2011). Thus, we see that examining actual behaviour on the social network is crucial to understanding the effects that these systems have on the well-being of users.

Most studies examining social capital outcomes from online social network usage have either examined overall usage of the site or application or the usage of specific features and interface elements. The second approach is exemplified by the question Yoder and Stutzman (2011) attempt to address : "where in the Facebook interface is social capital generated?". While usage statistics of various features and interface elements do constitute an aspect of online behaviour, we argue that research examining social capital in online social networks have largely overlooked another important aspect of social network behaviour : social network structure. Social network structure reveals implicit usage and interaction patterns in social networks, and its study has a long history in the traditional social sciences (Eg. Granovetter 1973; Burt 1995). Thus this source of

information can be valuable in broadening our current understanding of social capital in the context of online social networks.

## 2.3 Some Characteristics of the Problem Space

***Approaches to operationalizing Universal Behavioural Tendencies***

Broadly speaking, there are two ways in which universal behavioural tendencies have been considered while studying online behaviour. The first approach is to measure the various propensities of individuals in a population for a universal behavioural tendency. Once these propensities are measured, we can then correlate them with other measures of the individuals such as age, income or social capital. This approach can be considered as the application of differential psychology, the study of differences between individuals in terms of their propensity for various behavioural tendencies, and the consequences of these differences (Chamorro-Premuzic, 2011), to online behaviour. Literature adopting this approach suggests that differences between individuals in terms of personality traits and skills is related to how they use social networking systems (Anderson et al, 2012). For example, Ryan and Xenos (2011) investigated the influence of numerous personality constructs such as shyness, narcissistic tendencies, the big five personality trait factors and loneliness on the usage or non-usage of Facebook. Among their findings is that Facebook users are more likely to be extraverted and narcissistic, but they also have stronger feelings of family loneliness, and are more likely to be conscientious, shy, and socially lonely than non-users (Ryan and Xenos, 2011). In addition, demographic features such as gender also influence privacy attitudes (Stutzman & Kramer-Duffield, 2010). Thus overall we can expect that individual differences affect behaviour and usage in online social networks.

The second approach is to study how a particular universal behavioural tendency, like imitation, manifests in the population. In this approach we can ask questions such as "Under what contextual circumstances are individuals likely to imitate?", "What does imitation look like?" and "What are the effects of imitation on the bonding between individuals?". In relation to this approach, research looking at online social network behaviour can draw from a wealth of research in traditional social science. For example, social science suggests that individuals are likely to seek out and connect with those who are similar to them, in social background, interests, income, etc, a phenomenon

referred to as homophily (McPherson et al, 2001). This phenomenon has also been observed in online social networks, such as among the users of the Yahoo Instant messaging service (Aral et al, 2009), and among editors on Wikipedia (Crandall et al, 2008). In another study, Leskovek et al (2010) draw from social science theories such as Heider's theory of structural balance (Heider, 1958) to examine the interplay between positive (friendly) and negative (antagonistic) links in online social media.

We broadly refer to the first of the above approaches of operationalizing universal behavioural tendencies as the *individual differences* approach, and the second of these approaches as the *behavioural tendency centred* approach. In the work presented in this thesis, Chapter 3, and to an extent Chapter 4, adopt the behavioural tendency centred approach to understanding universal behavioural tendencies, while Chapters 5 and 6 adopt the individual differences approach.

#### *Behavioural Tendencies and Unconscious mental processes*

An important aspect of behavioural tendencies is that they are driven to a large extent by mental processes that are not conscious, or only partially conscious, to the individual. The fact is that most of us know very little of our automatic behaviour patterns (Cialdini, 1984). As an example, individuals can reveal evidence of implicit racism, but they are often not conscious of their racist tendencies (Gaertner and Dovidio, 1986). We can use the term "unconscious mind", or simply "the unconscious" (Jung, 1971), to refer to the unconscious mental processes that operate in an individual. An individual's unconscious mind is a pervasive and powerful influence over her higher mental processes such as judgements, decisions and the reasons for her behaviour (Bargh and Morsella, 2008). Thus while dealing with behavioural tendencies we must keep in mind that we are typically dealing with automatic behaviour patterns that for the most part are not consciously reflected upon, and therefore our experimental methods need to adapt accordingly.

While it is not in the scope of this thesis to unearth the various unconscious mental processes that drive any particular universal behavioural tendency that we study, it is important to understand this point as we attempt to measure universal behavioural tendencies in a study. This is because people are usually not very aware of, and not very well able to report on the true causes of their behaviour (Nisbett and Wilson, 1977). We argue that this is one of the reasons why numerous studies have found discrepancies

between users' reported privacy preferences and their actual behaviour in online social networks (Eg. Norberg et al, 2007; Reynolds et al, 2011). It is for the same reason we argue that in order to reliably glean information on universal behavioural tendencies, it is important that we use appropriate means, such as by directly observing behaviour in a controlled experiment, or by using questionnaire items that are carefully designed and validated to capture these universal behavioural tendencies, rather than ask users to directly report them.

#### *Online and Offline Behaviour*

A study of the relationship between online interactions and social capital is incomplete without considering offline interactions and behaviours. Indeed, one of the keys ways in which online social networks can be said to affect our well-being is by facilitating offline interactions. For example, online social networks allow us to co-ordinate offline events and meetings, or maintain latent offline ties. While there exist social networking sites and interactions in these sites that are purely between online ties, popular social networking sites such as Facebook are primarily concerned with people who already know each other, and use the Internet as one way of keeping their existing offline connections alive (Boyd and Ellison, 2007). Conversely, ties created online can turn into offline ties when the individuals decide to meet face-to-face. This is typical, for example, on the Couchsurfing (Lauterbach et al., 2009) social network.

Moreover, as explained in our conceptual map (Figure 1 in Chapter 1), both these behaviours are directly influenced by our behavioural tendencies. Thus, although the affordances of online and offline modalities are distinct, we should expect commonalities in the social networking behaviours in them as the behaviours in both modalities are influenced by a common driving force. Thus it is important to consider the interplay between online interactions and offline interactions if we are to understand how online interactions affect our well-being.

Given the close relationship between our online and offline social networks, it can be difficult to clearly draw a line between them in terms of the effects they separately have on reported outcomes such as social capital. Therefore, for the sake of understanding these modalities we make a distinction between them when we can, and where appropriate we clump them together and study outcomes as a result of a combination of these two modalities of interaction.

## 2.4 Research Questions

***Reciprocity and the Disclosure of Personal Information***

There has been a body of work examining various aspects of self-disclosure in the context of online social networks, such as the perceptions and interpretations of these disclosures (Bazarova 2012b), the strategies users adopt to control the visibility of these disclosures (Stutzman & Kramer-Duffield, 2010; Gonçalves et al., 2013), and their privacy attitudes towards such disclosures (Norberg et al., 2007; Reynolds et al., 2011). Our interest is to study the effect of universal behavioural tendencies on online behaviour, and particularly on self-disclosure. For our first study, we pick the universal behavioural tendency to reciprocate and examine the effect it has on self disclosure in online interactions. Reciprocity is an important universal behavioural tendency because it is one that every known human society has been reported to subscribe to (Gouldner, 1960). The norm arises fundamentally from the need for fairness and cooperation, and as human beings we feel obliged to return favours that we receive (Gouldner, 1960).

The reciprocity effect has been widely reported in the self-disclosure literature in social science and psychology (e.g. Archer and Berg, 1978; Collins and Miller, 1994). People seem to give back more, and more intimate information depending on the amount and kind of information received. Being a fundamental driver in human behaviour, we can ask whether reciprocity also plays an important role in online communication - are individuals likely to reciprocate the sharing of personal information in online interactions? Probing into how reciprocity operates in online disclosure can be important for two reasons. First, it can enhance our understanding of the dynamics of how people interact and how ties are formed and strengthened online. Second, given that personal information is central to privacy and security, understanding the relationship between reciprocity and information disclosure can help us get a grip on privacy predicaments and risks, especially those issues that are hard to tackle by technological innovations alone.

The first fundamental question of interest is whether *individuals reciprocate the disclosure of personal information in the online medium*. While significant differences

exist between the online medium of communication and face-to-face interactions, one might expect that due to the fundamental nature of this behavioural tendency, individuals might also tend to reciprocate the disclosure of personal information online. Further, online social networks pose an exception to the assumption that a large body of prior work, both on face-to-face and online interactions, is based on, which is that these disclosures take place in personal, one-to-one interactions.

That is, increasingly, online social networks facilitate public or 'broadcast' channels via which users disclose information in a one-to-many manner, not directed towards any particular individual. The profile page is an example of such a channel, where the user can disclose information about himself to a large audience. Another example is the 'wall post' where the user can broadcast information publicly or to an audience of friends. Thus, the second fundamental question of interest is whether *the reciprocity norm also holds when the initial disclosure is broadcast rather than directed to anyone in particular.*

In Chapter 3 we present a study that investigates the above two research questions in the context of the sharing of personally identifiable information with strangers in an online social network.

#### *Understanding Disclosure patterns through Network Structure*

Prior work has examined self-disclosures in online social networks (Eg. Bazrova 2012b; Stutzman & Kramer-Duffield, 2010) and also network structure in online social networks (Eg. Aral et al., 2009; Leskovek et al., 2010). However, the combined examination of these two proxies for online interaction behaviour is little explored. One reason for this is that, depending on the nature of the online social network being studied and the kind of information being exchanged between users in the social network, it can be a challenge to capture and measure instances of self-disclosures on a large scale. For example, if the users in the social network exchange information through text messages, we would require a computer program that can sift through these texts and detect instances of self-disclosure, designing which is a challenge in itself.

One way in which this difficulty can be bypassed is by using the privacy preference settings of each user to construct measures of how much each user shares, or is likely to share, with her various friends on the network. This can be especially fruitful when the

social networking system provides mechanisms that enables the user to control, for each of her friends, how much of the information she shares is visible to that friend. The location sharing social networking application Locaccino (Toch et al., 2010) is one such system. Locaccino has expressive rule creating mechanisms that allow users to define which of their friends can see their whereabouts at various times and places.

We take advantage of this feature of Locaccino to study the relationship between individuals' positions in the network and the disclosure of location information. We do so by reasoning based on prior literature to speculate how *various factors, including universal behavioural tendencies, might affect how different individuals share information with each other,* and then empirically test our speculations by analyzing the privacy policies of users with network structure. This work is presented in Chapter 4.

***Network Structure, Personality Traits and Social Capital***

There is a growing body of evidence that online social networks affect social capital (Eg. Best and Kruger, 2006; Burke et al, 2011; Yoder and Stutzman, 2011). These studies have typically analyzed various usage metrics such as time spent on social networking site, the features of the site that are used, which interface elements the user spent time on, and so on. Social science, on the other hand, has suggested another source of information that may be used to understand social capital - social network structure (Eg. Granovetter 1973; Burt 1995). There is little work in HCI that has attempted to take advantage of this source of information for understanding social capital. Thus, we take this insight from social science that social network structure is related to social capital and apply it to online social networks, to study the relationship between online social network structure and social capital.

We also attempt to study the effects of universal behavioural tendencies in the relationship between network structure and social capital. We do so by examining possible moderating influences of personality on the relationship between network structure and social capital. The motivation for doing this is that different individuals, depending on their personality traits, might differently tap into their network for social capital. For example, some individuals might be comfortable approaching a distant relative for help with employment, while others might not be comfortable doing so. This work is presented in Chapter 5.

### *Empathy and Social Network Structure*

We probe further into how universal behavioural tendencies affect the relationship between network structure and social capital by examining empathy in this context. Empathy can be described as the ability to feel or imagine another person's emotional experience (Lawrence et al., 2004), is fundamental to successful human relationships. Empathy allows us to understand the intentions of others, predict their behaviour, and experience an emotion triggered by their emotion (Baron-Cohen and Wheelwright, 2004). Given that empathy and social interactions are closely tied, *is empathy reflected in social network structure*? If so, then networks analysis can be used as a lens with which to study empathy, to the extent to which empathic skill is tied to social interactions. Motivated by this basic question, in this work we adopt a network science perspective to investigate how online social network structure can help us understand empathy and its relation to social capital. This is presented in Chapter 6.

## Chapter References


Amichai-Hamburger, Y. and Vinitzky, G. (2010). Social network use and personality. Computers in Human Behavior, 26(6):1289–1295.

Anderson, B., Fagan, P., Woodnutt, T., & Chamorro-Premuzic, T. (2012). Facebook psychology: Popular questions answered by research. Psychology of Popular Media Culture, 1(1), 23.

Aral, S., Muchnik, L., Sundararajan, A. (2009). Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. Proc Natl Acad Sci USA 106(51):21544–21549.

Archer, R.L. and Berg, J.H. (1978). Disclosure reciprocity and its limits: a reactance analysis. J. Exp. Soc. Psychol., 14, 527–540.

Bachrach, Y., Kosinski, M., Graepel, T., Kohli, P., & Stillwell, D. (2012). Personality and patterns of Facebook usage. In Proceedings of the 3rd Annual ACM Web Science Conference (pp. 24-32). ACM.

Barak, A. and Gluck-Ofri, O. (2007). Degree and reciprocity of self-disclosure in online forums. Cyber Psychol. Behav., 10.3, 407–417.

Bargh, J. A., Morsella, E. (2008). The unconscious mind. Perspect. Psychol. Sci. 3, 73–79 10.1111/j.1745-6916.2008.00064.

Bazarova, N. N. (2012a) Contents and Contexts: Disclosure Perceptions on Facebook. In Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work (CSCW '12). pp. 369–372. Seattle, Washington, WA, USA.

Bazarova, N. N. (2012b) Public intimacy: disclosure interpretation and social judgments on Facebook. J. Commun., 62, 815–832.

Best, S. J., and Krueger, B., S. (2006). Online Interactions and social capital : Distinguishing between new and existing ties. Social Science Computer Review 24(4), 395-410.

Boyd, D. M., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. Journal of Computer-Mediated Communication, 13(1), 210-230.

Burke, M., Marlow, C., and Lento, T. (2010). Social Network Activity and Social Well-Being. Proc. CHI 2010, ACM Press.

Burke, M., Kraut, R. & Marlow, C. (2011). Social capital on Facebook: Differentiating uses and users. Conference on Human Factors in Computing Systems CHI 2011, 571-580, ACM.

Burt, R. S. (1995). Structural Holes: The Social Structure of Competition. Harvard University Press.

Chamorro-Premuzic, T. (2011). Personality and individual differences (2nd ed.). Oxford, UK: Blackwell-Wiley.

Cialdini, R.B. (1984). Influence: The new psychology of modern persuasion. New York: Quill.

Collins, N.L. and Miller, L.C. (1994) Self-disclosure and liking: a meta-analytic review. Psychol. Bull., 116, 457–475.

Crandall, D., Cosley, D., Huttenlocher, D., Kleinberg, J., & Suri, S. (2008). Feedback effects between similarity and social influence in online communities. In Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 160-168). ACM.

Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The benefits of Facebook "friends:" Social capital and college students' use of online social network sites. Journal of Computer-Mediated Communication, 12(4), 1143-1168.

Gaertner, S. L., & J. F. Dovidio. (1986). The aversive form of racism. In J. F. Dovidio & S. L. Gaertner (Eds.), Prejudice, discrimination and racism: Theory and research (pp. 61–89). Orlando, FL: Academic Press.

Gilbert, E., & Karahalios, K. (2009). Predicting tie strength with social media. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 211-220). ACM.

Goncalves, J., Kostakos, V., & Venkatanathan, J. (2013). Narrowcasting in social media: effects and perceptions. In Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (pp. 502-509). ACM.

Gouldner, A.W. (1960) The norm of reciprocity: a preliminary statement. Am. Soc. Rev., 25, 161–178.

Granovetter. M. (1973). The strength of weak ties. American Journal of Sociology, 78:1360–1380.

Hampton, K. and Wellman, B. (2002). “The not so global village”, in B. Wellman and C. Haythornthwatie (eds.). The Internet in Everyday Life, pp 3-44, Oxford: Blackwell.

Heider, F. (1958). The Psychology of Interpersonal Relations. John Wiley & Sons.

Jung, C.G., (1971) . Psychological Types, Princeton University Press, Princeton, N.J., 1971. (Originally published in 1921.)

Kraut, R., Michael P., Vicki L., Sara K., Tridas M. and William S. (1998). “Internet Paradox: A Social Technology that Reduces Social Involvement and Psychological Well-Being?” American Psychologist 53 (9): 1017-1031.

Lampe, C., Ellison, N. B., & Steinfield, C. (2008). Changes in use and perception of Facebook. In Proceedings of the ACM 2008 conference on Computer supported cooperative work (pp. 721-730). New York: ACM.

Lauterbach, D., Truong, H., Shah, T., & Adamic, L. (2009). Surfing a web of trust: Reputation and reciprocity on couchsurfing. com. In Computational Science and Engineering, 2009. CSE'09. International Conference on (Vol. 4, pp. 346-353). IEEE.

Leskovec, J., Huttenlocher, D., & Kleinberg, J. (2010). Predicting positive and negative links in online social networks. In Proceedings of the 19th international conference on World wide web (pp. 641-650). ACM.

Meltzoff, A.N. and Moore, K.M. (1977) Imitation of facial and manual gestures by human neonates. Science, 198, 75–78.

Nie, N. H. (2001). Sociability, interpersonal relations, and the Internet. American Behavioral Scientist, 45, 420-435.

Nisbett R. E., Wilson T.D. (1977). Telling more than we can know: Verbal reports on mental processes. Psychological Review. 1977;84:231–259

Norberg, P., Horne, D., Horne, A.D.: The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors, Journal of Consumer Affairs, 41 (1), 100-126 (2007)

Papacharissi, Z., & Mendelson, A. (2010). Toward a new (er) sociability: uses, gratifications and social capital on Facebook. Media perspectives for the 21st century, 212.

Petronio, S. (2002). Boundaries of Privacy: Dialectics of Disclosure. SUNY, Albany, NY

Portes, A. (1998). Social Capital: Its Origins and Applications in Modern Sociology. Annual Review of Sociology, 24(1), 1–24. doi:10.1146/annurev.soc.24.1.1

Putnam, R. D. (2000). Bowling Alone: The Collapse and Revival of American Community. New York: Simon & Schuster.

Quan-Haase, A., & Wellman, B. (2004). How does the internet affect social capital? In M. Huysman & V. Wulf (Eds.), Social capital and information technology (pp. 113-135). Cambridge, MA: MIT Press.

Reynolds, B., Venkatanathan, J., Gonçalves, J., & Kostakos, V. (2011). Sharing ephemeral information in online social networks: privacy perceptions and behaviours. In Human-Computer Interaction–INTERACT 2011 (pp. 204-215). Springer Berlin Heidelberg.

Rubin, Z. (1975) Disclosing oneself to a stranger: reciprocity and its limits. J. Exp. Soc. Psychol., 11, 233–260.

Ryan, T., & Xenos, S. (2011). Who uses Facebook? An investigation into the relationship between the Big Five, shyness, narcissism, loneliness, and Facebook usage. Computers in Human Behavior, 27(5), 1658–1664. doi:10.1016/j.chb.2011.02.004

Stutzman, F. and Kramer-Duffield, J. (2010) Friends Only: Examining a Privacy-Enhancing Behavior in Facebook. Proc. Conf. Human Factors and Computing Systems: CHI 2010, pp. 1553–1562. ACM Press.

Steinfield, C., Ellison, N., Lampe, C., and Vitak, J. (2012). Online social network sites and the concept of social capital. In Lee, F. L., Leung, L., Qiu, J. S., and Chu, D. (eds.), Frontiers in New Media Research, New York: Routledge, 115-131.

Subrahmanyam, K., Reich, S.M., Waechter, N., & Espinoza, G. (2008). Online and offline social networks: Use of social networking sites by emerging adults. Journal of Applied Developmental Psychology 29 (6), 420-433.

The Wall Street Journal. (2012). Facebook: One Billion and Counting. Retrieved on September 3, 2014, from http://online.wsj.com/article/SB10000872396390443635404578036164027386112.html

Toch, E., Cranshaw, J., Hankes-Drielsma, P., Springfield, J., Kelley, P., Cranor, L., Hong, J. and Sadeh, N. (2010). Locaccino: A privacy-centric location sharing application. In Proceedings of the 12th ACM international conference adjunct papers on Ubiquitous computing-Adjunct (pp. 381-382). ACM.

Wellman, B. and Wortley, S. (1990). Different Strokes from Different Folks: Community Ties and Social Support. American Journal of Sociology 96, 3, 558-588.

Yoder, C., & Stutzman, F. (2011). Identifying social capital in the Facebook interface. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 585-588). ACM.

# 3. Online Disclosure and the tendency to Reciprocate

## 3.1 Introduction

In this chapter we present paper (a): "*Online disclosure of personally identifiable information with strangers: Effects of public and private sharing*", published in Interacting with Computers. The work explores the effect of the universal behavioural tendency to reciprocate, on how individuals share personal information online. While reciprocity is a universal behavioural tendency, it can also be seen as a social norm - it is an unwritten, unspoken "rule" that demands that we must give back, in fair value, what we have received from another. If someone lends you money, you should lend them money when they are in need. If someone gives you an expensive gift on your birthday, you should give them something expensive on their birthday. If someone invites you to dinner, you feel obliged to invite them to dinner sometime. The norm arises from the principle of fairness, and as human beings we feel obliged to return favours that we receive (Gouldner, 1960).

Reciprocity has also been widely reported in the self-disclosure literature (e.g. Archer and Berg, 1978; Collins and Miller, 1994). People seem to give back more, and more intimate information depending on the amount and kind of information received. Being a fundamental driver in human behaviour, we can ask whether reciprocity also plays an important role in online communication - are individuals likely to reciprocate the sharing of personal information in online interactions?

Therefore in this work, we examine the role that reciprocity plays in nudging people into sharing information about themselves in online interactions. Along with the theme of reciprocity, the paper combines the theme of personally identifiable information and how the human tendency to reciprocate can be exploited to conduct a privacy attack or identity theft. While the theme of reciprocity as a fundamental behavioural tendency is central to this thesis, the latter theme of personally identifiable information and privacy demonstrates a practical implication of the research.

## 3.2 Main Article - Online Disclosure of Personally Identifiable Information with Strangers: Effects of Public and Private Sharing

The article is organized in seven main sections. The first three sections introduce the research questions that we attempt to answer with our study and present the relevant background literature on which we can base our hypotheses. The fourth section describes in detail the study that we conducted in an online social network to understand the effect of the universal behavioural tendency of reciprocity on the disclosure of personally identifiable information. The fifth section presents the results of the study. Part of the analysis presented in section 5.1 contains minor oversights in the computation of the statistics. These do not change the interpretations drawn in the subsequent sections of article as statistical significance of the computed values remain unaffected. We direct the interested reader to the section of this thesis chapter following the article (Section 3.3), where we report the updated statistics and elaborate on its interpretation. The sixth section of the article discusses the results from our study, how it contributes to our understanding of the universal behavioural tendency to reciprocate in the context of online social networks, the implications for privacy and the relevance of the work in other areas of research in online social networks. The article ends with a discussion of the limitations of the study in final section.

*This is a pre-copy-editing, author-produced PDF of an article accepted for publication in Interacting With Computers following peer review. The definitive publisher-authenticated version [Jayant Venkatanathan et al. (2014). Online Disclosure of Personally Identifiable Information with Strangers: Effects of Public and Private Sharing Interact. Comput. 26(6), 614-626. doi:10.1093/iwc/iwt058] is available online at: http://iwc.oxfordjournals.org/content/early/2013/11/12/iwc.iwt058.full?ijkey=MVh2lVb1bsTWr8M*

# Online Disclosure of Personally Identifiable Information with Strangers: Effects of Public and Private Sharing

Jayant Venkatanathan[1,*], Vassilis Kostakos[2], Evangelos Karapanos[1] and Jorge Gonçalves[2]

[1] Madeira-ITI, University of Madeira.

[2]Department of Computer Science and Engineering, University of Oulu.

[*]*Corresponding Author: jayant.venkatanathan@m-iti.org*

**Abstract.** Safeguarding personally identifiable information (PII) is crucial because such information is increasingly used to engineer privacy attacks, identity thefts and security breaches. But is it likely that individuals may choose to just share this information with strangers? This study examines how reciprocation can lead to the disclosure of PII between strangers in online social networking. We demonstrate that the widespread use of public, one-to-many communication channels such as 'wall posts' and profile pages in online social networks poses an exception to the assumption that reciprocation happens on one-to-one channels. We find that individuals not only reciprocate and share PII when the disclosure of such information is private and directed towards them by a stranger, but also when the stranger shares this information through a public channel that is not directed towards anyone in particular. Implications for privacy and design are discussed.

## 1. Introduction

Individuals are increasingly turning to online social networks to draw support in their day-to-day activities and pursuits. These sites range from support groups for smoking cessation (Cobb *et al*., 2010) and weight loss (Hwang *et al*., 2010) to travel and accommodation (Lauterbach *et al*., 2009) and language learning (Harrison and Thomas, 2009). Individuals often tend to disclose information about themselves to each other in these interaction settings. Indeed, mutual disclosure of personal information facilitates the development of trust and bonding between individuals. However, such disclosures can also be potentially drawn and exploited by malicious parties attempting to carry out a social engineering attack.

With the increasing adoption of online social networks, and the increasing sophistication of social engineering attacks, an important research challenge is to develop an understanding of how social mechanisms and norms can be exploited for potential attacks. Such an understanding is crucial for the designers of social

networking sites in order to foresee these potential attacks and put design mechanisms in place to prevent them.

In this paper, we focus on a social norm that is crucial for the understanding of information disclosure in a social network setting—reciprocity. The paper demonstrates that reciprocity is an important factor that can leads to the disclosure of personally identifiable information (PII) in online social networks. While there exists substantial evidence showing that individuals tend to reciprocate the act of sharing information about themselves when in one-to-one situations (e.g. Archer and Berg, 1978; Barak and Gluck-Ofri, 2007), it is not clear whether this holds for online social networks where communication patterns are also one-to-many. Increasingly, however, online social networks facilitate public or 'broadcast' channels via which users disclose information in a one-to-many manner, not directed towards any particular individual. The profile page is an example of such a channel, where the user can disclose information about himself to a large audience. Another example is the 'wall post' where the user can broadcast information publicly or to an audience of friends. Such widespread use of broadcast communication channels in online social networks poses an exception to the assumption that a large body of prior work both in the context of online and face-to-face interactions is based on, which is that these disclosures are made in personal, one-to-one interactions.

Furthermore, while previous work has focused on 'self- disclosure', researchers tend to group together a broad range of information about oneself ranging from inner feelings and thoughts (of fear, vulnerability, etc.) to more mundane and factual information. In the context of online social networks, the reciprocity of PII is of particular interest because it can be used to engineer a privacy attack, identity theft or security breach. Hence, we are interested in the reciprocation of PII in the context of online social network interaction with strangers.

In examining whether a reciprocity norm exists in the disclosure of PII (such as full name, occupation, date of birth, nationality, etc.) in the online space, one might expect that the type of channel (public one-to-many vs. private one-to-one) through which the first person discloses her details can influence the other person's decision to reciprocate that disclosure. Hence, the two main research questions that the study reported in this paper addresses are: (1) Is there a reciprocity norm for the disclosure of PII in the online space? and (2) Does the initial disclosure of such information have to be one-to-one in order for the reciprocity norm to come into effect?

The rest of this paper is organized as follows: We first present an overview of prior work in the fields of social engineering attacks and self-disclosure, leading to the two fundamental research questions outlined above (Sections 2 and 3). We then present a study that we designed and conducted in an online social network in order to answer these research questions. Through the study, we demonstrate

how the norm of reciprocity can be exploited by malicious parties to draw PII from unsuspecting users (Sections 4 and 5). We go on to discuss how these results enhance our prior understanding of reciprocation in online social networks and how it can be exploited for social engineering attacks. We discuss the implications for the design of social networking systems that can foresee and prevent such attacks with appropriate mechanisms in place (Section 6). Finally, we outline the limitations of this work and the ground that it sets for future work to explore (Section 7).

# 2. Background

## 2.1. Phishing, social engineering attacks and PII

A deliberate intrusion into, by unjust means, or exploitation of an individual's information or credentials in an online context is referred to as an attack. As a hypothetical example, in a privacy attack an employer might covertly intrude into a potential employee's online social network in order to draw information for a 'background check', thus causing a violation of her privacy. The term attack can also be used while referring to the methodology used to carry out the intrusion, such as a phishing attack or a social engineering attack. A phishing attack is one in which the attacker attempts to con a victim into divulging personal information using spoofed emails and fraudulent websites. Rather than exploiting bugs in computer software, in a phishing attack the attackers attempt to directly extract sensitive information from a victim by posing as a legitimate source (Downs *et al*., 2007). Direct phishing-related losses to US financial institutions have been estimated at over a billion dollars per year (Emigh, 2005). Thus, phishing poses a significant challenge to online security.

A particularly effective form of phishing, known as spear phishing or *social engineering attacks*, involves personalized messages incorporating elements of context (O'Brien, 2005). Literature on phishing suggests that users are aware that they need to protect their computer from problems like malware, but are less aware of social engineering attacks aimed at eliciting information directly from them (Downs *et al*., 2006).

It has been suggested that as phishers get smarter, future generations of phishing attacks will incorporate more elements of context to become more effective (Jagatic *et al*., 2007). For an attacker to incorporate these elements of context in an attack, an important first step would be to obtain a user's PII. PII can be defined as information that can be used to distinguish or trace an individual's identity either alone or when combined with other information that is linkable to a specific individual (Krishnamurthy and Wills, 2009). This includes information that can by itself uniquely trace an individual's identity such as complete name, social security number or biometric records, or in combination with other data such as date of birth or mother's maiden name (Johnson, 2007).

PII research has shown that individual pieces of personal information, when linked together from different sources, can be surprisingly accurate in identifying an individual. For example, a study by Acquisti and Gross (2009) demonstrates that people's social security numbers can be predicted based on other pieces of data such as birth date and birth location. Another well-known result in linking pieces of PII is that most Americans (87%) can be uniquely identified from a birth date, five-digit zip code and gender (Malin, 2005).

Thus, safeguarding PII is crucial because such information can be used to engineer privacy attacks and identity thefts (Moyer and Hamiel, 2008). Moreover, certain websites such as banks require users to enter their date of birth along with account information such as the credit card number as a fallback authentication mechanism when they forget their password (Rabkin, 2008). Hence, such PII can be potentially misused by malicious parties to gain access to users' accounts. Therefore, an important research challenge is to develop an understanding of how social mechanisms can be potentially exploited for attacks that attempt to extract PII from unsuspecting users. With such an understanding, we can foresee these potential attacks and enable the designers of these systems to put mechanisms in place that prevent them.

## 2.2. Online disclosure and reciprocity

There is a significant body of work on understanding the disclosure of a wide array of information about oneself, ranging from biographical data to more intimate information such as opinions, beliefs and fears (e.g. Archer and Berg, 1978; Collins and Miller, 1994). This research has mostly been clumped together under the term 'self-disclosure'. Self-disclosure has been defined as any personal information that a person communicates to another (Altman and Taylor, 1973; Collins and Miller, 1994) and it builds trust by making the discloser increasingly vulnerable to the other person (Rubin, 1975). Altman and Taylor (1973) categorize self-disclosure into three layers: peripheral (biographical data, age, etc.), intermediate (attitudes, values, opinions, etc.) and core (personal, beliefs, needs, fears and values).

The reciprocity effect (Gouldner, 1960) has been widely reported in the self-disclosure literature (e.g. Archer and Berg, 1978; Collins and Miller, 1994). People seem to give back more, and more intimate information depending on the amount and kind of information received. Further, it has been shown that people who disclose more tend to be liked more and people disclose more to those they initially like (Collins and Miller, 1994). However, the nature of the relationship between individuals is an important factor. For example, the obligation to reciprocate disclosure may be stronger between strangers than between friends (Derlega and Chaikin, 1975, p.50). Self- disclosure has been used as a tactical means to elicit information, such as in police interrogation of suspects (Alison *et al.*, 2007).

Online disclosure may not involve certain vulnerabilities associated with offline disclosures, due to the relative anonymity and the ability to control which matters one wishes to reveal (Ben-Ze'ev, 2003). Hence, people seem to disclose more intimate information in Internet relationships (Parks and Floyd, 1996). Joinson and Paine (2007) remark that the relationship between self-disclosure and privacy is paradoxical—privacy is a prerequisite for disclosure, yet the process of disclosure serves to reduce privacy. On examination of prior work such as the above, we can infer that the lack of PII (which is implied by anonymity) facilitates the disclosure of more subjective information such as fear, desire and personal shortcomings in online interactions. Hence, PII is distinctly different from more subjective personal information when it comes to individuals' needs to share such information. Yet, to our knowledge, prior work on self-disclosure and the reciprocity of self-disclosure has largely failed to make an explicit separation of PII in examining the disclosure of information about oneself. This type of disclosure is, effectively, a disclosure of identity.

There is also a body of recent work in HCI examining the extent to which individuals disclose personal information, and the methods and strategies adopted by them to manage these disclosures. An early study of Facebook showed that the majority of users disclosed PII on their profile pages (Gross and Acquisti, 2005). In addition, there is often a discrepancy between people's privacy attitudes towards sharing information and their actual sharing patterns (Acquisti and Gross, 2006; Norberg *et al*., 2007, Reynolds *et al*., 2011). This behaviour has been termed the 'privacy paradox'. For instance, a study revealed a high discrepancy between stated concerns and actual behaviour towards sharing static profile information on Facebook (Acquisti and Gross, 2006). Privacy regulation in social networking sites can be considered a socio- technical activity involving interaction with the technological system and the group context. Individuals' privacy behaviour in such systems involves a mixture of technical and mental strategies. For instance, a technical strategy may involve the use of privacy settings to regulate content distribution to select audiences, such as only friends in the system (Stutzman and Kramer-Duffield, 2010), while research has also shown that considering tie strength can be another strategy for developing rules for disclosure (Wellman and Wortley, 1990). Complex group dynamics also play a role in how individuals share information. For example, individuals who occupy more central positions, in terms of the structure of the social network, tend to reveal more information (Kostakos *et al*., 2011).

A large part of the work in HCI such as the above are in social network sites that are primarily concerned with people who already know each other, and use the Internet as one way of keeping their existing social connections alive (Boyd and Ellison, 2007). While this is not surprising given that social networks such as Facebook are among the most accessed websites, there exist other important and popular social networking services in which, due to their nature and purpose,

interactions can occur between strangers and often between different cultures and regions (Harrison and Thomas, 2009). These contexts of online social interaction have been largely unexplored in the HCI literature. The online social network Livemocha, with over 9 million users as of the time of writing this manuscript, is an example of such a social network where interactions are typically between strangers. Thus, what can be termed as self-disclosure in such a context can be very different from that on sites such as Facebook, as self-disclosure is not merely characterized by the information that is shared, but also by the context of the interaction (Antaki *et al*., 2005). Moreover, sites such as Livemocha, unlike Facebook, have relatively rudimentary mechanisms for managing the level of exposure, ruling out privacy management strategies such as restricting access to only friends (Stutzman and Kramer-Duffield, 2010) or 'narrowcasting' each post only to the audience for which it is intended (Goncalves *et al*., 2013).

The reciprocity effect in 'self-disclosure' has been previously reported in online media. For example, one study observed such reciprocity in online forums where 'self-disclosure' was measured by adding together instances of disclosures of facts, thoughts and feelings about oneself (Barak and Gluck-Ofri, 2007). There has also been recent work on understanding public disclosures. For example, on Facebook, disclosure shared privately is perceived to be more intimate than disclosure shared publicly (Bazarova, 2012a,b). However, no work to our knowledge has specifically examined the reciprocity of PII disclosure, and such reciprocity in the context of broadcast disclosures.

The study described next examines reciprocity in the context of disclosing PII to strangers in online social networks. More specifically, it examines the effect of reciprocity in the disclosure of one's full name and date of birth with strangers in an online social network, both in a one-to-one and one-to-many context.

# 3. Hypotheses

Previous work suggests that people tend to reciprocate the act of disclosing a broad range of information about themselves (Joinson and Paine, 2007). Thus, in the context of online social networks, individuals may be more likely to disclose PII if they do so in reciprocation. This reasoning provides grounds for the first experimental hypothesis:

> H1: Individuals are more likely to reveal PII with a stranger in an online social network when reciprocating.

While previous work has reported the reciprocity effect with respect to a range of social exchanges where the initial disclosure is personal and one-to-one, this does not provide us with any grounds to hypothesize whether reciprocity can come to play when the initial disclosure is public and one-to-many. In other words, if a stranger posts his full name and date of birth on his public profile page, and then requests from another user her full name and date of birth, does this bring into

play a norm of reciprocity that makes the user more likely to reveal this information? There are no clear grounds for us to suspect that such a request is as likely to result in compliance as in the case in which the stranger shares personal identification in a one-to-one message directed to the target user. This leads us to the second hypothesis:

> H2: Individuals given PII in a one-to-many interaction are less likely to reveal this information than those who are given this information in a one-to-one interaction.

In other words, H2 hypothesizes that such a reciprocity norm only holds in situations where the initial disclosure is one-to-one and directed to an individual.

# 4. Method

It is methodologically challenging to capture behaviours of users with regard to disclosure of PII in technology-mediated interactions, in a *realistic* manner and setting. Previous work has identified a discrepancy between people's attitudes and stated preference towards sharing information and their actual behaviour (Acquisti and Gross, 2006; Norberg *et al*., 2007). Thus, in order to preserve the authenticity of the setting and the validity of our results, we adopted a method to directly observe users' behaviour, as followed by (Jagatic *et al*., 2007).

Asking participants for informed consent would nullify our experiment. Thus, we opted to obtain implicit consent by giving participants an opportunity to respond (or not) to messages we sent them, and then fully debriefed and rewarded all participants at the end of the study. All participants were rewarded within the context of the social network we study, a community-based language learning website, by offering them help in language learning and providing feedback on their language exercises. An alternative approach would be to ask potential participants for informed consent for a fictional study, and then introduce our experimental stimulus. We felt this was inappropriate in our study because it may affect our results due to participants being suspicious, while at the same time it would involve lying to participants in a public online setting that could impose further stress on them.

## 4.1 Study design overview

We designed a study in order to test our two hypotheses. The study involved sending a message from an experimental profile to individuals in an online social network, attempting to elicit their full name and date of birth. The online social network chosen, Livemocha, is one in which interactions are typically between strangers. Owing to this, these pieces of information were to an extent privacy sensitive in the context of the social network. These target individuals from whom we attempted to elicit information were allocated to one of three conditions, and the condition determined the manner in which we attempted to elicit this

information. In condition A, the experimental profile did not divulge his own full name or date of birth in his messages. This was the control condition. In condition B, the experimental profile divulged his own full name and date of birth in his messages. Hence, condition B served to test hypothesis H1. In condition C, the experimental profile did not divulge his own full name or date of birth in his private messages, but had posted asd

this information on his public profile page. Hence, condition C was used to test hypothesis H2.

The study was conducted on Livemocha, an online social network for language learning, with over 9 million registered users as of the time of writing this manuscript. For each language listed on the website, there are written exercises that involve writing a small paragraph in that language. A user learning a particular language can complete these exercises, and users who are speakers of that language can provide feedback on these exercises. To encourage participation, the website allows users to become 'friends', send private messages and chat with each other.

Each Livemocha user has a profile page where they can upload a profile picture, write a description and share other details about themselves such as age and location. Most people choose to upload a profile picture. Since most of the social interaction is initiated around the submission and correction of exercises, interactions on Livemocha are often between individuals from different cultures or countries, who typically have not met each other before. All profiles are visible to all users, and there are no detailed privacy mechanisms to obscure parts of one's profile to certain individuals.

Compared with Facebook, Livemocha is a much more 'low- tech' website. It lacks the dynamic interface elements found on Facebook, does not have rich media capabilities or search capabilities, and is particularly tuned to one purpose: learning languages. The benefit of this approach is that profile informa- tion and privacy settings are very explicit and easy to under- stand, unlike in Facebook where users often complain about not being able to understand who can see their information.

While Livemocha does not have complicated privacy mechanisms, like in Facebook, it does have certain mechanisms to help users determine credibility. For each profile, one can see the date of registration, indicating whether a user has just registered or is a seasoned veteran. In addition, users get points as a reward for being active on the site. For instance, users are awarded points when correcting an exercise submitted by another user. The total points are also visible for each profile, thus indicating the extent to which a user is a 'good citizen' on the site.

## 4.2 Study Procedure

Our first step was to crawl 26 000 randomly selected, publicly available profiles on Livemocha, using the profiles' unique identifier as the random seed. Analysis of these data indicated to us that the most popular native language on the website was Portuguese. This led us to decide to target Brazilian Portuguese speakers, as their large presence was expected to speed up data collection. In addition, we found that, for every English speaker learning Portuguese, there were 22 Portuguese speakers learning English. This mismatch between Portuguese and English speakers suggested that if our experimental profiles spoke English, then users with complementary skills and needs are most likely to respond.

Following this initial analysis, we next created experimental profiles in Livemocha that were listed as Indian males who were English speakers. Details such as gender and nationality were identical across all the experimental profiles in order to keep results comparable between them. Each experimental profile submitted a beginner-level written exercise in Brazilian Portuguese that consisted of two simple sentences with a few simple grammatical errors. We designed the exercise submission, with the help of native speakers, to be extremely easy to correct, in order to minimize the effort that the participants would invest in our study. Subsequently, we waited for speakers of Brazilian Portuguese to provide feedback on this exercise.

Once Livemocha users responded to the exercises submitted by our experimental profiles, we sent messages from the respective experimental profile to those users, attempting to elicit their full name and date of birth. This request was made under the pretext of interest in understanding their culture (Table 1). Once users responded to this message, they were briefed about the study being conducted, via a profile belonging to one of the researchers.

The study ran between February and June 2011. A total of 10 experimental profiles were created (4 for condition C and 6 for conditions A and B together). Participants were allocated to condition A or B based on the alternating order of time at which they provided feedback to the exercise submitted by the experimental profiles used for these conditions. Since condition C required the experimental profile to have additional information in the profile page, the experimental profiles used in this condition were minimally different from those used in conditions A and B (Fig. 1). Each experimental profile was used only once to submit an exercise and subsequently message those users who provided feedback to the exercise. This was done to keep to a minimum the 'activity' level of experimental profiles, as that can introduce changes across profiles. Therefore, all experimental profiles were newly registered and had uniformly low credibility in terms of user points and teacher points awarded by the Livemocha system automatically.

A total of 99 participants provided feedback to the exercises submitted by the experimental profiles and each participant was subsequently messaged. One participant provided feedback to the exercise of two experimental profiles, and

these data were discarded. The total participants were 35 (12 male) in condition A, 34 (18 male) in condition B and 30 (11 male) in condition C.

For each participant, we recorded: age, the date of joining Livemocha, gender, 'user points' and 'teacher points' as reported by Livemocha. The user points reflect the extent of the total activity of the user on the website which includes completing lessons, submitting exercises and submitting feedback on other users' exercises. The teacher points the extent of the user's activity on the website in terms of the feedback he or she has provided on others' exercises. Age and gender were optional data that the participants could fill in. Seventy-four out of the 99 participants listed their age (mean 29.25, s.d. 11.5, median 26.5) on their profile page.

A total of 59 (28 male) participants responded to the message from the experimental profile. We refer to these participants as 'respondents'. Forty-three of these respondents had listed their age (mean 28.7, s.d. 10.25, median 26).

| Messages used in Condition B | Messages used in Conditions A & C |
|---|---|
| *Thanks a lot for correcting my exercise. I am very interested in learning about Brazilian culture. In my town in India people use their father's name as their surname. So my full name is "Ashok Mohan" where Ashok is my name and Mohan is my father's name. How is it in Brazil? There are some amusing things about Indian culture. For example, I was born on 2 November in 1982, and I am considered lucky because it was the birth anniversary of a god named Krishna. When were you born? Is your date of birth special in any way?* | *Thanks a lot for correcting my exercise. I am very interested in learning about Brazilian culture. In my town in India people use their father's name as their surname. So if you saw a name like "Vinay Mohan", Vinay is the guy's name and Mohan is actually his father's name. How is it in Brazil? There are some amusing things about Indian culture. For example, I have a friend who was born on 2 November in 1982, and he is considered lucky because it was the birth anniversary of a god named Krishna. When were you born? Is your date of birth special in any way?* |

**Table 1. Messages that were used in the three conditions**

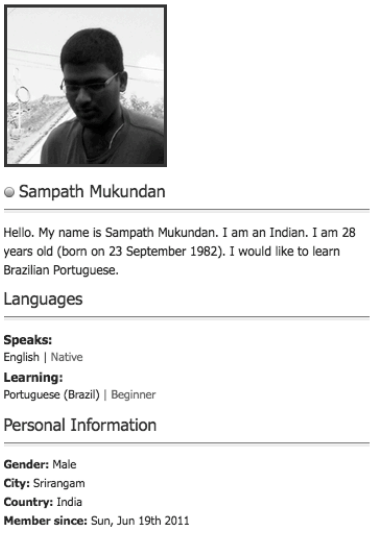


**Figure 1. Screenshot of the information provided on the profile page of an experimental profile used in condition C. Experimental profiles used for conditions A and B were identical except that the description field (with full name and date of birth) was blank.**

# 5. Results

A binary logistic regression test examining the decision to reply or not to the bait message (sent by the experimental profile) did not result in significance for any of the variables recorded: condition, age, gender, user points, teacher score ($P > 0.05$).

Subsequently, we analysed the effect of various variables on whether respondents revealed information pertaining to both kinds of PII that the experimental profile attempted to elicit from them, i.e. name and date of birth. More specifically, we consider that a participant has disclosed his full name if he mentions it in his message to the experimental profile and this mentioned name consists at least of two distinct parts (i.e. the participant mentions a first name and a last name).

When it comes to date of birth information, certain participants disclosed their birthday to the experimental profile, while certain participants, in addition to their birthday, also mentioned their year of birth. On the other hand, the year of birth of many participants could easily be inferred, given the large number of participants who mentioned their age on their profile page (75 out of 99). Therefore, it is not clear whether those who mentioned only their birthday did so with an intention to hide their year of birth or did so because it was not relevant to the significance of birth dates in Brazilian culture. Hence, for date of birth information, we consider whether a participant disclosed their birthday (not accounting for whether they revealed their year of birth) to the experimental profile.

## 5.1. Effect of different variables on disclosure

Participants from conditions B and C (52 and 50%, respectively) were equally likely to reveal their full name and birthday, followed by those in condition A (15%) (Fig. 2). The complete results are summarized in Table 2.

We conducted a hierarchical logistic regression to analyse the effect of various variables on whether respondents revealed both their full name and birthday. For our response variable in the regression, we gave 'Revealed full name and

| | Condition A *No disclosure* | Condition B *One-to-One* | Condition C *One-to-Many* |
|---|---|---|---|
| Total Users | 35 | 34 | 30 |
| Respondents | 20 | 23 | 16 |
| Revealed Full Name | 4 | 16 | 10 |
| Revealed Birthday | 8 | 15 | 9 |
| Revealed Full Name AND Birthday | 3 | 12 | 8 |
| Revealed Full Name AND DOB, year of birth in message | 1 | 8 | 3 |

**Table 2. Summary of information revealed by users across three conditions**

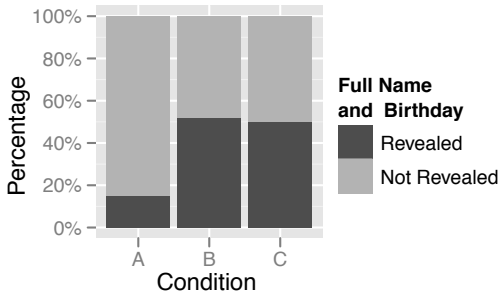


**Figure 2. Respondents in conditions B and C were about equally likely to reveal PII, while those in condition A were less likely to do so.**

birthday' responses a value of 1 and 'Did not reveal full name and birthday' a value of 0. Our primary objective was to examine the effect of our experimental manipulation, i.e. the differences between the conditions. Therefore, for our main explanatory variable we used the condition to which respondents belonged.

In addition, we also incorporated the age, gender, time on the website, user points and teacher points as explanatory variables. Incrementally adding blocks of variables to the model allowed us to examine whether the newly incorporated variables provided improved prediction ability over the preceding model. However, given our sample size, we must interpret the results pertaining to these additional variables with caution. The primary objective and contribution of this work is to examine the effect of the experimental manipulation across the three conditions, and further variables are examined only to draw additional insights into disclosure patterns.

Table 3 shows the parameters for the logistic regression and our resulting analytical model of sharing decisions. In the first stage of our model, we examined if the condition in which the participants were allocated affected their decision to reveal their full name and birthday. The results showed that the condition to which the participants belonged offered significant predictive power to our model ($P < 0.01$). The model also included a significant negative constant (intercept) component ($B = -1.73$, $P < 0.01$), indicating that by default our

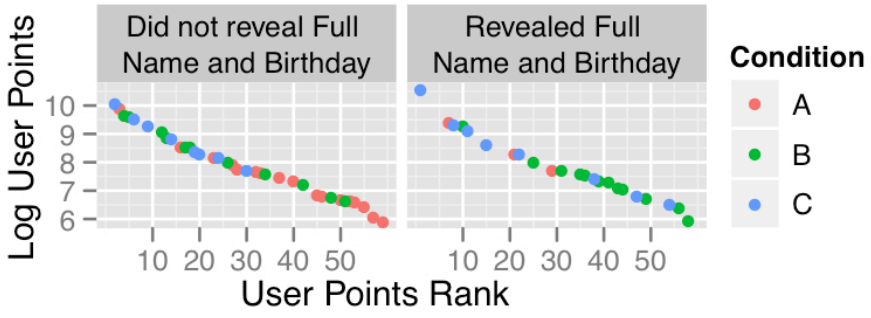


**Figure 3. User points had no effect on whether respondents revealed PII or not. On the x-axis is the rank of user points among respondents. The points are spread out across the entire range among both the group of users who revealed this information and those who did not.**

| | | B(SE) | z-value | P(>\|z\|) | exp(B) |
|---|---|---|---|---|---|
| STEP 1 | Condition(1) | 2.41(1.12) | 2.15 | 0.032 | 11.13 |
| | Condition(2) | 2.89(1.16) | 2.50 | 0.014 | 17.99 |
| | Intercept | -2.89(1.03) | -2.81 | 0.005 | 0.056 |
| STEP 2 | Condition(1) | 2.59(1.24) | 2.01 | 0.037 | 13.33 |
| | Condition(2) | 3.18(1.29) | 2.46 | 0.014 | 24.05 |
| | Gender | -1.06(0.84) | -1.25 | 0.210 | 0.346 |
| | Age | 0.049(0.040) | 1.29 | 0.219 | 1.05 |
| STEP 3 | Condition(1) | 1.76(0.80) | 2.19 | 0.028 | 5.81 |
| | Condition(2) | 2.66(0.97) | 2.73 | 0.006 | 14.30 |
| | Time on Website | 0.000(0.000) | 1.145 | 0.147 | 1.00 |
| | User Points | 0.000(0.000) | -0.652 | 0.515 | 1.00 |

**Table 3. Details of binary logistic regression modelling the factors involved in the prediction of sharing decisions.**

participants did not exhibit an inclination to reveal their full name and birthday unless other variables motivated them to do so. A likelihood ratio test showed that the improvement of this model over the null model was statistically significant ($\chi 2(2) = 8.43, P < 0.05$).

In the second stage of our model, we examined if participants' demographics could account for any variation in their choice to reveal their full name and birthday. We found that age and gender did not offer significant influence ($P > 0.05$) within our model, and were hence removed from the subsequent stage.

In the third stage of our model, we examined if respondents' experience of using the website affected their decision to reveal their full name and birthday. The teacher points variable was not included in this equation as it had a high correlation with the user points variable (Pearson's correlation = 0.88, $P < 0.001$). The results show that the time since people registered on the website or their user points did not significantly affect their decision to reveal this information ($P > 0.05$). Figure 3 also illustrates that user points had no effect.

Since we did show the age of experimental profiles used in condition C on the profile page, we checked for the effect of this on participants' disclosure. To avoid suspicion, we had varied the age reported on the profile page of these experimental profiles. We reported the age of the four experimental profiles used in this condition as 28, 28, 29 and 21, respectively. The experimental profile with age 21 had the highest rate of respondents who disclosed this information (4 out of 4), but the ages of these respondents greatly varied (17, 35, 60 and 33). However, the effect of the experimental profiles' age on whether participants revealed their full name and birthday was not significant ($\chi^2$ (6) = 8.0, P = 0.24).

Profiles in condition B too revealed their age in the private message, but this was constant in all messages sent by experimental profiles in this condition (28 years).

We also examined the effect of various variables on whether participants mentioned their full name and complete date of birth, including year of birth, in their message to the experimental profile. A binary logistic regression showed no statistically significant difference between conditions A and C in this analysis ($P > 0.05$), although participants in condition B were significantly more likely to mention this information ($P < 0.05$). As in the analysis of the disclosure of full name and birthday (not accounting for year of birth), gender, age or experience on the website did not have any significant effect on the likelihood of participants revealing this information.

## 5.2. Qualitative information in responses and follow-up questions

We received responses in both English and Portuguese across participants. Some participants mentioned that they had used an online translation tool since they were not good at English. One respondent from condition A asked for the birthday of the experimental profile.

> *I was born on <date removed> . . . How about you? When is your birthday?*

Some respondents did not divulge their own name but gave examples instead. Instances of these were found in all three conditions. For example, one participant wrote

> *Let's imagine my mother's name is <name removed>, and the name of my father is <name removed> . . . the child's name might look like <name removed>.*

Some participants merely explained how the full name is derived from the mother's and father's names without giving an example. Some responses to mentioning their name and date of birth were brief and to the point, while others were elaborate.

While the text in most messages pertained to the explanation of names and the significance of dates of birth, some respondents divulged other details such as interests and employment.

Overall, responses from conditions B and C tended to be longer (mean 130 words (s.d. 90) and mean 115 words (s.d. 72), respectively) than those from condition A (mean 102 words (s.d. 61)).

At the end of the study, all participants (including those who did not respond to the message of the experimental profile) were informed about the study being conducted. We apologized for needing to have communicated with them through an experimental profile, and explained why it was necessary for us to have done

that in order to observe responses in a valid manner. As a gesture of appreciation, we offered them help with their English exercises. We were interested in understanding better their behaviour with the experimental profile, and seven users offered give us further feedback through an optional questionnaire. Five of these users had responded to the experimental profile's message while two had not. From this feedback, we learnt that all were active users of other social networks such as Facebook and Orkut, and some had used Internet banking and made online transactions. Thus, this subset of participants were to an extent seasoned users of the Internet.

Overall these participants initially felt that there was some genuineness in the experimental profile's interest in Brazilian culture. They found it interesting for an outsider to be interested in their culture, and wanted to help such a person in learning about it. When asked why they did or did not share their full name or date of birth with the experimental profile, one of the users wrote that she was tricked by the 'complete casualness' of the message into sharing her details. Finally, those who shared any information with the experimental profile reported to have shared accurate information.

# 6. Discussion

Our results show that users were much more likely to reveal their full name and date of birth when the experimental profile revealed his own. This suggests that individuals tend to reciprocate the act of sharing PII (more specifically full name and date of birth information), confirming hypothesis H1.

On the other hand, individuals who could see the full name and date of birth information of the experimental profile on his public profile page were more likely to reveal their information than those who were not given this information. Since condition C was identical to A in terms of the message received by the user, the only factor that can explain the significant difference in the disclosures in this condition is that these users subsequently went to the profile page of the experimental profile and saw the additional information posted there. As a result of seeing additional details posted on the profile page, these users were more willing to share their details.

Moreover, there was no difference between conditions B and C when it came to disclosure of full name and birthday, leading us to reject H2. That is, participants were equally willing to reveal this information irrespective of whether the experimental profile shared his information in a private message or in a broadcast manner. This provides evidence that the reciprocity norm implied by H1 also applies to the case where the initial disclosure is one-to-many.

## 6.1. The norm of reciprocity

This paper set out to address two fundamental questions with regard to the sharing of PII with strangers in an online social network. The first is whether individuals reciprocate the sharing of such information. Our results indicate that the answer to this question is yes. This is in agreement with prior work on 'self-disclosure' taken as a disclosure of a broad range of personal information (e.g. Barak and Gluck-Ofri, 2007).

More surprising, however, is the finding that the reciprocation occurs even when the information is broadcast, such as through a public profile page, where it is not directed at a particular user. This is especially interesting in the light of recent findings that public disclosures on Facebook were perceived less intimate than private disclosures (Bazarova, 2012a,b). Our findings suggest that in stranger interactions, there might be no difference between public and private disclosures of personal identifiable information in terms of willingness to reciprocate such disclosures.

It must be noted that the users who were sent a message had all first provided feedback on an exercise submitted by the experimental profile. This was done in order to increase the rate of response to the messages. In addition, the fact that users might have perceived the experimental profile to be able to help them with learning English might have increased response rates overall. Consequently, the reciprocation that we have observed is over and above these effects. However, since these factors apply equally to all three conditions, the conclusions drawn from our results remain valid.

The simplest interpretation of our results is that the sharing of the full name and date of birth affected the compliance of the recipient when it came to revealing his own full name and date of birth because the recipient felt obligated to reciprocate the act. By sharing PII, an individual communicates a certain degree of trust on the recipient, and it is an unspoken obligation of the recipient to reciprocate this act of trust when required to do so (Derlega and Chaikin, 1975). Hence, the reciprocity of disclosing PII can also be viewed fundamentally as a reciprocity of a display of trust. Interestingly, this display of trust can be towards a group or community of people as a whole and the norm of reciprocity still holds when an explicit request is made to an individual from this group.

## 6.2. Effects beyond reciprocity

Even though a reciprocity norm is a plausible explanation for the increased compliance observed in our results, one cannot rule out other causes. We next discuss possible alternative explanations for the results we have obtained, and show whether or not they are plausible. While the following list is not meant to be exhaustive, we believe these are the most important alternative factors that can explain the observed results.

*Credibility*: It can be hypothesized that the act of revealing their full name and date of birth made the experimental profile seem more credible. Therefore, using credibility as a guiding concern (e.g. Andrade *et al*., 2002), respondents showed increased compliance in conditions B and C. However, we argue that if credibility was indeed the guiding concern, all conditions would have observed a low level of compliance. This is due to the fact that their credibility was actually quite low: all profiles were newly created, with extremely low user points and teacher points, indicating a person who is not an active or trusted member of the community. Hence, we can rule out credibility as the guiding concern of respondents, as they all responded to overall low-credibility profiles.

*Imitation*: Studies have shown that humans have a tendency to imitate the behaviour of others (e.g. Meltzoff and Moore, 1977). Along similar lines, it is plausible that respondents in conditions A and B tended to replicate the behaviour of the experimental profile in their response by hiding or disclosing their full name and date of birth in their message. However, this hypothesis does not account for the behaviour of respondents in condition C. If these respondents were simply imitating, then they should not be more likely to disclose their details than respondents in condition A, since the message they received was identical in both conditions. On the other hand, imitation might partially explain why participants in condition B were more likely to explicitly mention their year of birth in the message, since the experimental profile mentioned his year of birth in the message too. Participants in condition C, however, were possibly less disposed to do so, as year of birth was probably irrelevant to explaining the significance of their birthday, and the experimental profile himself did not mention his own year of birth in his message. Nevertheless, while imitation might possibly explain the difference between conditions B and C in terms of disclosure of year of birth, it does not fully explain our results.

*Liking*: Research has shown a link between 'self-disclosure' and liking, which can in turn lead to self-disclosure in return (Collins and Miller, 1994). Here, the motivation for disclosure is not a feeling of obligation. Rather, this explanation posits that because in conditions B and C the experimental profiles shared personal information, respondents felt that they like this profile. As a result, they chose to also share their personal details. While we cannot completely rule out this hypothesis, there is evidence against it. Primarily, all profiles were mostly identical: the nationality, gender and approximate age of the experimental profiles were identical, therefore equally contributing to a respondent's liking of the profile. It is true that in conditions B and C the profile shared a date of birth, which may have had an impact on respondents' liking of the profile. While we cannot rule it out, this explanation asserts that the reciprocity effect we have

observed is indirect. In either case, the impact of our experimental manipulation is existent, whether direct or indirect.

*Erroneous norm-inference:* Visibility of actions allows individuals to observe others' behaviour and infer social norms (Erickson and Kellog, 2000). Thus, the experimental profile publicly revealing his personal information might have suggested to new users that sharing such information is a norm on the website. As a result, respondents in condition C might have been more willing to share this information. However, the data suggest that this is not the case. First, our observations showed that sharing such private information is in fact not a norm on this website. However, one might expect that new users may not be aware of this, and could potentially be 'tricked' into believing this behaviour is a norm. It is also possible that technology savviness and prior experience with the web may have a role to play in this. While we do not have data for technology savviness or overall web usage in our sample, our analysis of experience on the website (time since registration, teacher points, user points) showed no relationship with whether users shared their information. While we caution the reader to interpret with care the results from variables in addition to our experimental conditions, at least within our sample respondents who shared their information were at various levels of experience on the website. This can also be visually verified by Fig. 3— the dots representing users in condition C appear across the range of user point values. Hence, this explanation is unlikely to explain our results.

## 6.3. Implications for privacy

The experiment described in this paper demonstrates the vulnerability of users against attempts to trick them into revealing information by exploiting this social norm. Inferring or linking personal information such as that obtained in the current study would typically be an important first step in a malicious party's attempt to exploit a user. For example, the malicious party might use elements of context inferred from the site such as the user's interest in learning a language or the people that the individual has friended in the social network. The unsuspecting user might then be sent a spam message or email incorporating these elements of context on his birthday for an advertisement of a language learning product or even a link to a virus. Such context-aware spam messages are known to have higher click-through rates (Brown *et al*., 2008) and are thus likelier to trick the user.

With people increasingly interacting with strangers on various social networking platforms, there is a need for mechanisms to help them identify such attackers apart from genuine users. Exploiting social norms and trust is a well-understood mechanism for social engineering attacks (Jagatic *et al*., 2007). What our results show, however, is that whereas such attacks were targeted in a one-on-one fashion, users are also vulnerable to easier and cheaper one-to-many attacks.

While systems must support the development of ties between individuals, and mutual disclosure of personal information is an integral element of such a bonding process, it is important to distinguish between genuine individuals forming a relationship and malicious parties. The challenge is then to provide mechanisms that help users identify such malicious parties in a manner that does not hinder the sharing of information between genuine users.

In looking for a solution to this problem, we might take inspiration from the theory of social translucence (Erickson and Kellog, 2000). The authors of that work suggest that making certain activity or information visible ('translucence') to relevant individuals can cause those directly involved to act differently. It does so through supporting *mutual awareness* among all individuals ('they know that others know') and this brings our social rules into play and therefore a sense of accountability on the part of those who are acting. Clearly, there is a tension between visibility and privacy, and the goal is not to take away the privacy of the environment but rather to understand that privacy simply supports certain types of behaviour and inhibits others. Drawing from this idea, one solution to protect users from such attacks would be to provide a public communication channel for each profile, similar to the Facebook 'wall'. This allows an individual who is approached by a stranger attempting to elicit private information (under a pretext such as interest in culture, as in our study) to move their interaction to this public space where she can address his supposed interest without divulging in personal information. This provides a certain amount of visibility of the users' interactions to the community. When other users view the stranger's wall, they know that the stranger has been doing this with multiple users and thereby exercise caution in their interactions with him. While such a solution does not eliminate the risk of this kind of attack, it serves as a means for users to support each other and reduce its chances.

Another approach would be to automatically monitor newly created profiles and profiles that have not invested much effort in the activities of the community (in our case, users with low teacher and user points). If such a user sends messages with similar content to multiple recipients within a short time span, the low credibility level of the user can be highlighted to the recipients, so that they can make an informed decision to exercise caution. However, such a solution must be thoughtfully implemented, as it might result in disproportionate costs for genuine newcomers and thus the community as a whole, since newcomers are crucial for the vitality of online communities (Kraut *et al*., 2010). It is therefore important to keep perspective of genuine forms of interaction so as to ensure that the solution does not inhibit them.

## 6.4. Relevance to Facebook research

Unlike the most popular social networks such as Facebook that are better explored in HCI, where ties are mostly between individuals who share some

offline element (Boyd and Ellison, 2007), in the social network examined in this study interactions are typically between strangers. In Facebook, users connect with people from different aspects of their lives, including family, friends, schoolmates and co-workers. Thus, the issue of context collapse (Boyd, 2008)—the process by which various kinds of individuals' ties become grouped together under generic terms such as 'Friends'—and how users manage the merging of these different contexts is important to understand in sites such as Facebook.

One might argue that in sites such as Livemocha the issue of context collapse is relatively less complex, as users in the social network do not share information with people with different aspects from their lives but rather connect with people for the purpose of language learning and to explore cultures. Nevertheless, reciprocity can be an important factor in information disclosure on Facebook as is on sites such as Livemocha. For example, if a user on Facebook shares her phone number with her friend on a wall post or comment that is visible to a large audience, is her friend likely to feel compelled to share her phone number too on the same thread? While work has examined how users perceive and interpret disclosures (Bazarova, 2012a,b), it would also be interesting to study how they perceive *non-disclosures* such as the refusal to directly reciprocate. For example, if the above friend does not share her phone number in the wall post, or possibly chooses to rather share her number in a private message, how would the user and the audience to which the wall post is visible perceive this? It would be interesting and fruitful for future work to explore how the processes of reciprocity and context collapse operate together, and perhaps contrast this between online and face-to- face social networks (Kostakos and Venkatanathan, 2010).

# 7. Limitations and Future Work

A methodological drawback of the study is not that participants were chosen at random from the population of users but rather that participants were self-selected. Therefore, we cannot rule out the presence of a non-response bias in our sample, whereby those users who chose not to correct our exercise or reply to our message might have behaved differently from our observed sample. While the implications for privacy remain unchanged, the extent of reciprocity observed in our results might possibly differ from that of a truly random sample. However, this methodology is more valid in our case than those that rely on self-reported data from users. For example, it might be unrealistic to expect to obtain credible data by asking users questions such as 'Would you reveal your complete name and date of birth to a stranger in an online social network?'.

On the other hand, our own and others' recent work in understanding social engineering attacks (e.g. Jagatic *et al*., 2007) has resorted to using a post-consent technique, to directly observe users. While such naturalistic experiments must be executed with caution and avoided where possible, there is an important case for

them in understanding online fraud (Jakobsson *et al*., 2008). With the increasing sophistication of social engineering attacks (Jagatic *et al*., 2007) it might be important for researchers to develop and test alternative lab methods, such as role play, to understand online fraud. Nevertheless, because the approach used in this paper offers the most accurate picture, the need to better understand how people interact with computers makes such research worthwhile (Jakobsson *et al*., 2008).

While discussing privacy aspects, a valid question to ask is whether the personal information used in the study is privacy sensitive within the context of which these disclosures took place. One can argue that information such as full name is easily accessible nowadays, as opposed to a credit card number or a social security number, and hence users' perceptions of privacy threats might have been low. On the other hand, the Livemocha social network is largely anonymous where users have never met before and typically share no mutual friends or other social support mechanisms that they can use to socially verify each other's credibility. In such a context, one might expect users to build trust over time and multiple interactions. Hence, it is of concern that such information can be elicited within one or two brief interactions, as was observed in this study. While the findings on reciprocity hold irrespective of the extent to which such information is privacy sensitive within the current context, future work must explore how users behave when it comes to more sensitive information and how far one can take this before the reciprocity effect breaks down.

We also highlight that for the majority of participants we did not verify whether the date of birth they reported was correct. Some users deliberately use fake personal details online to minimize their exposure to fraud, and it is possible that some respondents adopted this strategy when responding to our bait message. It is certainly interesting to investigate further the extent of this behaviour and its consequences.

Finally, we note that our experimental profiles were all Indian males who listed themselves as English speakers and our sample consisted only of native Brazilians. This was an explicit decision we made to make our experimental profiles more attractive, as these profiles could provide help in English learning. This is likely to have resulted in a potential power imbalance, which might have had an effect on the participants' willingness to respond. Future work can examine the effect of such power imbalances in the context of information disclosure. Further, although it is expected that the reciprocity observed in this study also holds for a more general population of users, clearly there might be differences across cultures in finer details, such as the extent to which such a norm is adhered to. These potential cultural differences set a fertile ground for future work to explore.

# Acknowledgements

The authors would like to thank Filipe Quintal and Lucas Pereira for their support over the period of the study. This work was funded by the Portuguese Foundation for Science and Technology (FCT) grant CMU-PT/SE/0028/2008 (Web Security and Privacy). Additional support was provided by the Academy of Finland and TEKES.

# References

Acquisti, A. and Gross, R. (2006) Imagined Communities: Awareness, Information Sharing, and Privacy on the Facebook. In Proceedings of the 6th international conference on Privacy Enhancing Technologies (PET'06), pp. 36–56. Springer-Verlag, Berlin, Heidelberg.

Acquisti, A. and Gross, R. (2009) Predicting Social Security Numbers from Public Data. Proc. Natl Acad. Sci. USA, 106, 10975.

Alison, L., Kebbel, M. and Leung, J. (2007) Facet analysis of police officers' self-reported use of suspect interviewing strategies and their discomfort with ambiguity. Appl. Cogn. Psychol., 22, 468– 481.

Altman, I. and Taylor, D. (1973) Social Penetration: The Development of Interpersonal Relationships. Holt, Rinehart, & Winston, New York.

Andrade, E.B., Kaltcheva, V. and Weitz, B. (2002) Self-disclosure on the web: the impact of privacy policy, reward, and company reputation. Adv. Consum. Res., 29, 350–353.

Antaki, C., Barnes, R. and Leudar, I. (2005) Self-disclosure as a situated interactional practice. Br. J. Soc. Psychol., 44, 181–199.

Archer, R.L. and Berg, J.H. (1978) Disclosure reciprocity and it's limits: a reactance analysis. J. Exp. Soc. Psychol., 14, 527–540.

Barak, A. and Gluck-Ofri, O. (2007) Degree and reciprocity of self- disclosure in online forums. Cyber Psychol. Behav., 10.3, 407–417

Bazarova, N.N. (2012a) Contents and Contexts: Disclosure Percep- tions on Facebook. In Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work (CSCW '12). pp. 369–372. Seattle, Washington, WA, USA.

Bazarova, N.N. (2012b) Public intimacy: disclosure interpretation and social judgments on Facebook. J. Commun., 62, 815–832.

Ben-Ze'ev, A. (2003) Privacy, emotional closeness, and openness in cyberspace. Comput. Hum. Behav., 19, 451–467.

Boyd, D. (2008) Taken Out of Context: American Teen Sociality in Networked Publics. School of Information. University of California, Berkeley, Berkeley, CA.

Boyd, D. and Ellison, N. (2007) Social networks: definition, history, and scholarship. J. Comput.-Mediat. Commun., 13, 210–230.

Brown, G., Howe, T., Ihbe, M., Prakash, A. and Borders, K. (2008) Social Networks and Context-Aware Spam. In Proceedings of the 2008 ACM conference on Computer supported cooperative work (CSCW '08), pp. 403–412. ACM, San Diego, CA, USA.

Cobb, N.K., Graham, A.L. and Abrams, D.B. (2010) Social network structure of a large online community for smoking cessation. Am. J. Public Health., 100, 1282–1289.

Collins, N.L. and Miller, L.C. (1994) Self-disclosure and liking: a meta-analytic review. Psychol. Bull., 116, 457–475.

Derlega, V.J. and Chaikin, A.L. (1975) Sharing Intimacy. What we Reveal to Others and Why? Prentice-Hall, Inc., Englewood Cliffs, NJ.

Downs, J.S., Holbrook, M.B. and Cranor, L.F. (2006) Decision Strategies and Susceptibility to Phishing. In Proceedings of the second symposium on Usable privacy and security (SOUPS '06), pp. 79–90. Pittsburgh, PA, USA.

Downs, J.S., Holbrook, M.B. and Cranor, L.F. (2007) Behavioral response to phishing risk. In APWG 2nd Annual eCrime Researchers Summit, pp. 37–44. Pittsburgh, Pennsylvania, PA, USA.

Downs, J.S., Holbrook, M., & Cranor, L.F. (2007). Behavioral response to phishing risk. In Proceedings of the anti-phishing working groups 2nd annual eCrime researchers summit, pp. 37–44. ACM, Pittsburgh, PA, USA.

Emigh, A. (2005) Online Identity Theft: Phishing Technology, Choke- points and Countermeasures. Identity Theft Technology Coun- cil Report. http://www.antiphishing.org/Phishing-dhs-report.pdf (accessed October 3, 2005).

Erickson, T. and Kellog, W. (2000) Social translucence: an approach to designing systems that support social processes. Trans. Comput. Hum. Interact., 7, 59–83.

Goncalves, J., Kostakos, V. and Venkatanathan, J. (2013) Nar- rowcasting in Social Media: Effects and Perceptions. Proc. of ASONAM'13, Niagara Falls, Canada. IEEE.

Gouldner, A.W. (1960) The norm of reciprocity: a preliminary statement. Am. Soc. Rev., 25, 161–178.

Gross, R. and Acquisti, A. (2005) Information Revelation and Privacy in Online Social Networks. Workshop on Privacy in the Electronic Society, Alexandria, VA. ACM Press.

Harrison, R. and Thomas, M. (2009). Identity in online communities: social networking sites and language learning. Int. J. Emerg. Technol. Soc., 7, 109–124.

Hwang, K.O., Ottenbacher, A.J., Green, A.P., Cannon-Diehl, M.R., Richardson, O., Bernstam, E.V. and Thomas, E.J. (2010) Social support in an Internet weight loss community. Int. J. Med. Inform., 79, 5–13.

Jagatic, T.N., Johnson, N.A., Jakobsson, M. and Menczer, F. (2007) Social phishing. Commun. ACM 50, 10 (October 2007), 94–100.

Jakobsson, M., Finn, P. and Johnson, N. (2008) Why and how to perform fraud experiments. IEEE Secur. Priv., 6, 66–68.

Johnson III, C. Safeguarding Against and Responding to the Breach of Personally Identifiable Information, May 22 2007. Office of Management and Budget Memorandum.

Joinson, A.N. and Paine, C. (2007) Self-disclosure, Privacy and the Internet. In Joinson, A.N., McKenna, K.Y.A., Postmes, T. and Reips, U. (eds), Oxford Handbook of Internet Psychology, pp. 237–252. Oxford University Press, New York.

Kostakos, V. and Venkatanathan, J. (2010) Making Friends in Life and Online: Equivalence, Micro-Correlation and Value in Spatial and Transpatial Social Networks. SOCIALCOM '10, pp. 587–594. IEEE.

Kostakos, V., Venkatanathan, J., Reynolds, B., Sadeh, N., Toch, E., Shaikh, S.A. and Jones, S. (2011) Who's Your Best Friend? Targeted Privacy Attacks in Location-sharing Social Networks. UbiComp '11: Ubicomp, pp. 177–186. ACM, Beijing, China.

Kraut, R., Burke, M. and Riedl, J. (2010) Dealing with New Comers. In Kraut, R.E. and Resnick, P. (eds), Evidencebased Social Design Mining the Social Sciences to Build Online Communities: 1 42. MIT Press.

Krishnamurthy, B. and Wills, C.E. (2009) On the Leakage of Personally Identifiable Information Via Online Social Net- works. ACM SIGCOMM Workshop on Online Social Networks (WOSN).

Lauterbach, D., Truong, H., Shah, T. and Adamic, L. (2009) Surfing a Web of Trust: Reputation and Reciprocity on CouchSurfing. com. 2009 Int. Conf. Computational Science and Engineering, pp. 346–353. IEEE.

Malin, B. (2005) Betrayed by my shadow: learning data identify via trail matching. J. Priv. Technol.

Meltzoff, A.N. and Moore, K.M. (1977) Imitation of facial and manual gestures by human neonates. Science, 198, 75–78.

Moon, Y. (1998) Impression management in computer-based interviews: the effects of input modality, output modality, and distance. Public Opin. Quart., 62.

Moyer, S. and Hamiel, N. (2008) Satan is on my friends list: attack- ing social networks. http://www.blackhat.com/html/bh-usa-08/ bh-usa-08-archive.html.

Norberg, P.A., Daniel, R.H. and David, A.H. (2007) The privacy paradox: personal information disclosure intentions versus behaviors. J. Consum. Aff., 41, 100–126.

O'Brien, T.L. (2005) Gone spear-phishin'. The New York Times (4 December). http://www.nytimes.com/2005/12/04/business/yourmoney/04spear.html?pagewanted=1&ei=5088&en=2f313fc4b55b47bf&ex=1291352400&partner=rssnyt&emc=rss.

Parks, M.R. and Floyd, K. (1996) Making friends in cyberspace. J. Commun., 46, 80–97.

Reynolds, B., Venkatanathan, J., Goncalves, J. and Kostakos, V. (2011) Sharing Ephemeral Information in Online Social Networks: Privacy Perceptions and Behaviours. INTERACT, pp. 204–215. Springer, Berlin.

Rubin, Z. (1975) Disclosing oneself to a stranger: reciprocity and its limits. J. Exp. Soc. Psychol., 11, 233–260.

Stutzman, F. and Kramer-Duffield, J. (2010) Friends Only: Examining a Privacy-Enhancing Behavior in Facebook. Proc. Conf. Human Factors and Computing Systems: CHI 2010, pp. 1553–1562. ACM Press.

Wellman, B. and Wortley, S. (1990) Different strokes from different folks: community ties and social support. Am. J. Sociol., 96, 558–588.

## 3.3 Critical Review of the work and conclusion

***Revisiting the logistic regression analysis***

We summarise here the logistic regression analysis presented in Table 3 of the paper (Section 5.1). A few minor clarifications have been made, which are insignificant in terms of the effect they have on our subsequent discussion - none of the conclusions drawn from the statistics are affected. We nevertheless report the analysis here as it provides an accurate picture of the magnitude of association of the variables with the outcome of whether participants shared their personal information or not. In addition, we also elaborate on the procedure of the analysis and the interpretation of the resulting statistics to make it easier for the reader to comprehend the results.

Table 3.1 shows the parameters and resulting statistics for our logistic regression analysis. For our outcome variable we gave 'Revealed full name and birthday' responses a value of 1 and 'Did not reveal full name and birthday' a value of 0. The analysis is performed in a step-wise fashion, each step examining a new model by

| | | B (SE) | Wald's z-score | P(>\|z\|) | exp(B) |
|---|---|---|---|---|---|
| STEP 1 | Condition(B) | 1.82 (0.75) | 2.42 | 0.016 | 6.18 |
| | Condition(C) | 1.73 (0.80) | 2.17 | 0.030 | 5.67 |
| | Intercept | -1.73 (0.63) | -2.77 | 0.006 | 0.18 |
| STEP 2 | Condition(B) | 3.05 (1.20) | 2.56 | 0.011 | 21.21 |
| | Condition(C) | 3.31 (1.26) | 2.62 | 0.009 | 27.46 |
| | Gender(Male) | -0.80 (0.79) | -1.02 | 0.307 | 0.45 |
| | Age | 0.03 (0.04) | 0.70 | 0.481 | 1.03 |
| | Intercept | -3.00 (1.50) | -2.00 | 0.046 | 0.05 |
| STEP 3 | Condition(B) | 1.78 (0.77) | 2.33 | 0.020 | 5.94 |
| | Condition(C) | 1.86 (0.84) | 2.20 | 0.028 | 6.39 |
| | Time on Website (rescaled /100) | -0.04 (0.11) | -0.42 | 0.676 | 0.96 |
| | User Points (rescaled /10000) | -0.03 (0.05) | -0.56 | 0.576 | 0.97 |
| | Intercept | 0.15 (4.41) | 0.04 | 0.972 | 1.17 |

Table 3.1 Details of binary logistic regression modelling the factors involved in the prediction of sharing decisions.

adding blocks of variables to the preceding model. This allowed us to examine whether the newly incorporated variables provided improved prediction over the preceding model. Our approach here in testing for the significance of coefficients of variables in each step is along the lines of the following question : Does the model that includes the variables in question tell us more about the outcome variable than the model that does not include that variables? If the predicted values with the variables in the model are better than when the variables are not in the model, then we can consider the variables in question to be "significant".

The first column of Table 3.1 contains the estimates of the coefficients (B) along with their estimated standard errors (SE). Three additional columns are presented. One displays the Wald's z-score, the ratio of the estimated coefficient to the estimated standard error (B/SE). The next column displays the p-value of the Wald's z-score, which indicates the significance of the coefficient. The last column, exp(B), indicates the odds ratio, which provides a meaningful interpretation of the association of the predicting variable with the outcome variable.

Since the Wald's z-score reported in the tables provide us with a measure of the significance of the variables to the model, we could simply combine all the variables in a single step and examine the significance of each variable using the Wald score. However, inferences based on the Wald score are reported to be inadequate, sometimes failing to reject the null hypothesis when the coefficients are significant (Hauck and Donner, 1977). To overcome this inadequacy we need to use the likelihood ratio test, which is a more reliable test for the significance of predictor variables (See Hosmer and Lemeshow, 2004 for a discussion of this issue). This is why we choose to conduct our analysis in a step-wise manner by incrementally adding blocks of variables, and use the likelihood ratio test to determine whether the addition of the new variables improve the model.

In the first stage of our model we examined if the condition in which the participants were affected their decision to reveal their full name and date of birth or not. The Wald's scores suggested that the condition to which the participants belonged offered significant predictive power to our model ($p<.05$). Further, a likelihood ratio test showed that the improvement of this model over the null model was statistically significant ($\chi^2(2) = 7.97$, $P < 0.05$). The column exp(B) indicates the odds ratio for the

predictor variable, which is the change in the odds in favour of an outcome of 1 ("Revealed full name and birthday") with a unit change in the predictor variable. These results in step 1 indicate that the odds that a subject would share the full name and birthday for among those in conditions B and C are about 6 times (6.18 for condition B and 5.67 for condition C) the odds among subjects in condition A.

In the second step we examined if respondents' demographics could account for any variation in their choice to reveal their full name and birthday. As shown in the table, the Wald's scores suggest that the Gender and Age variables are not significant in the model (P>0.05). Moreover, a likelihood ratio test showed no improvement of this model over the model in Step 1 ($\chi^2(2) = 1.31275$, $p = 0.52$). Hence we discard these variables for the subsequent step.

In the third step we examined if respondents' experience of using the website affected their decision to reveal their full name and birthday. We used two variables, the time since people registered on the website, and their user points on the website, as measures of their experience of using the website. The difference in number of days between the earliest registered user and the latest registered user in our sample was 1,165 days, and the difference between the highest and lowest values of user points was 38,662 points. Since these are large ranges in comparison with those of the other variables in our analysis, we rescale them by dividing the time on the website by 100 and dividing the user points by 10,000. Rescaling does not affect the analyses but makes the statistics more easily interpretable. The Wald scores shown in the Table suggest that the time since people registered on the website or their user points are not significant variables in the model (P>0.05). Further, a likelihood ratio test showed no improvement of this model over the model in Step 1 ($\chi^2(2) = 0.3772$, $p = 0.83$). Thus we reject this model too, and retain the model in Step 1.

***Trust and Contextual Priming***

A final point that we would like to touch upon in this study is the issue of trust, and its role in the reciprocation of disclosures. Clearly, trust is central to disclosure of personal information - if a person does not feel sufficient trust in a given situation or context, she is less likely to disclose personal information, and therefore is also less likely to reciprocate the disclosure of personal information. While the behavioural tendency to reciprocate is likely to provide an additional nudge towards a disclosure, the contextual

conditions can support or inhibit the playing out of the behavioural tendency. In a given situation within an online social network, the contextual conditions from the perspective of a user include all visible factors, ranging from the people she is interacting with, the people who can see the interaction and the kinds of activities that the social network facilitates, to the design of various interface elements on the website.

This is akin to the phenomenon of "contextual priming" in physical spaces - in the presence of certain contextual conditions, in a physical space, people are more likely to play out certain behaviour patterns. For example, ambient stimuli (e.g. hammers) automatically set us to physically interact with the world (e.g. perform a power grip) (Tucker & Ellis, 2001). In a similar manner, we can view different online social networks, and even different "places" within an online social network (such as different group pages on Facebook), as spaces that provide different contextual conditions. Therefore people may be more likely to act out certain behavioural patterns in certain online spaces than in others.

This is likely to have been the case in our study too. For example, it is likely that our respondents, possibly due to having seen members help each other in language learning, or having themselves helped or received help on Livemocha, viewed the social network as a friendly and supportive space. As a result they might have been contextually primed to trust in that space, and were therefore willing to a certain extent to reciprocate the sharing of personal information with a stranger in that space. In section 6.2 of the article, we suggested that "credibility" did not seem to be the driving concern in respondents' sharing decisions, where we used the term "credibility" to refer to the reputation of profiles within the social networking site given by the system through user points. This means that despite the low reputation scores of our experimental profiles, the respondents were willing to share their information, perhaps due to contextual conditions as argued here. On the other hand, it may be possible to prime users to consider the reputation score in their process of trusting, should the designers of the system choose to do so. This might be done, for instance, by making the reputation score more salient and visible on the user interface, or by educating users to exercise some caution in their interactions with profiles that are not yet of good reputation within the community.

The issue of context and contextual priming also brings up the question of whether these results in terms of reciprocity are directly transferable to other online social networks and contexts. Obviously, it would be naive to say that we will observe disclosure reciprocity behaviour, in the same manner we observed in the context of this study, in every other kind of online setting and with every kind of personal information. We can only infer from this study that the tendency to reciprocate does provide an impulse to share in online - both one to one and one to many - contexts. This impulse is likely to be triggered across multiple online contexts, given the fundamental nature of the tendency to reciprocate. However, whether that impulse is strong enough to result in a disclosure or not, given the forces of other contextual factors, can only be gauged based on the particular context.

#### *Conclusion*

We set out in this chapter to study the effect of the universal behavioural tendency of reciprocity on online disclosure behaviour. We found evidence that individuals reciprocate the disclosure of personally identifiable information in the online medium. What is even more fascinating is that this tendency to reciprocate is triggered even when the initial disclosure is broadcasted and not directed to them in particular. This is a result one might not have foreseen only from understanding reciprocity in the context of face to face interactions. We discuss the implications of the results from the standpoint of online privacy and trust.

## Chapter References

The following are the references cited in this chapter outside of the embedded journal article.

Archer, R.L. and Berg, J.H. (1978). Disclosure reciprocity and its limits: a reactance analysis. Journal of Experimental Social Psychology, 14, 527–540.

Collins, N.L. and Miller, L.C. (1994). Self-disclosure and liking: a meta-analytic review. Psychol. Bull., 116, 457–475.

Gouldner, A.W. (1960). The norm of reciprocity: a preliminary statement. Am. Soc. Rev., 25, 161– 178.

Hauck, W.W. and Donner, A. (1977). Wald's test as applied to hypotheses in logit analysis. Journal of the American Statistical Association, 72:851–853.

Hosmer Jr, D. W. and Lemeshow, S. (2004). Applied logistic regression. John Wiley & Sons.

Tucker, M., & Ellis, R. (2001). The potentiation of grasp types during visual object categorization. Visual Cognition, 8, 769– 800.

# 4. Studying the Interplay between Disclosure Patterns and Network Structure

## 4.1 Introduction

In this chapter we present paper (b): "*Who's Your Best Friend? Targeted Privacy Attacks In Location-sharing Social Networks"* presented at the Ubicomp conference in 2011. This work brings together disclosure patterns and social network structure in order to study these two proxies for online interaction together. Prior work has examined self-disclosures in online social networks (Eg. Bazarova 2012; Stutzman & Kramer-Duffield, 2010) and also network structure in online social networks (Eg. Aral et al., 2009; Leskovek et al., 2010). However, the combined examination of these two proxies for online interaction behaviour is little explored, perhaps due to the fact that it is difficult to capture and measure instances of self-disclosures on a large scale.

One way in which this difficulty of scale, in terms of disclosure information, can be bypassed is by using the privacy preference settings of each user to construct measures of how much each user shares, or is likely to share, with her various friends on the network. This can be especially fruitful when the social networking system provides mechanisms that enables the user to control, for each of her friends, how much of the information that she shares is visible to that friend. It is increasingly the case in today's online social networking systems that they incorporate such control mechanisms as they evolve. For example, both Facebook and Google Plus provide mechanisms that attempt to make it convenient for users to pick out a subset of friends to share a particular post or piece of information. Thus the method we suggest to study disclosure information and network structure in conjunction can be potentially applied across a number of social networking systems.

In this work, we demonstrate the potential of this method by actually carrying out an analysis of disclosure patterns and network structure in conjunction, based on data from the location sharing social networking application Locaccino (Toch et al., 2010). Locaccino has expressive rule creating mechanisms that allow users to define which of their friends can see their whereabouts at various times and places. We take advantage of this feature of Locaccino to study the relationship between individuals' positions in the network and the disclosure of location information. We do so by reasoning based on

prior literature to speculate how various factors, including universal behavioural tendencies, might affect how different individuals share information with each other, and then empirically test our speculations by analyzing the privacy policies of users with network structure.

### 4.2 Main Article : Who's Your Best Friend? Targeted Privacy Attacks In Location-sharing Social Networks

The article motivates the work from the perspective of understanding privacy attacks, and therefore the first two sections (Introduction and Related Work) set up the theme of the paper from that perspective. In the third section (Study) we describe in detail the study we conducted based on data from the Locaccino real time location sharing system. This includes the definitions of the various metrics we use in our analysis and the procedure by which we derive those metrics. The next section (Results) describes the analysis of the data using these metrics. This section is followed by a discussion of the results from this analysis of the Locaccino data and a conclusion.

*Kostakos, V., Venkatanathan, J., Reynolds, B., Sadeh, N., Toch, E., Shaikh, S. A., Jones, S. (2011). Who's your best friend? Targeted privacy attacks in location-sharing social networks. In proceedings of Ubicomp 2011, Beijing, China, p. 177-186.*
 *doi: 10.1145/2030112.2030138*

# Who's Your Best Friend? Targeted Privacy Attacks In Location-sharing Social Networks

**Vassilis Kostakos, Jayant Venkatanathan, Bernardo Reynolds**
Madeira Interactive Technologies Institute
University of Madeira
{vk, vjayant, bernardo.reynolds}@m-iti.org

**Norman Sadeh, Eran Toch**
School of Computer Science
Carnegie Mellon University
{sadeh, eran}@cs.cmu.edu

**Siraj A. Shaikh**
Faculty of Engineering and Computing
Coventry University
s.shaikh@coventry.ac.uk

**Simon Jones**
Department of Computer Science
University of Bath
s.jones2@bath.ac.uk

**ABSTRACT**
This paper presents a study that aims to answer two important questions related to targeted location-sharing privacy attacks: (1) given a group of users and their social graph, is it possible to predict which among them is likely to reveal most about their whereabouts, and (2) given a user, is it possible to predict which among her friends knows most about her whereabouts. To answer these questions we analyse the privacy policies of users of a real-time location sharing application, in which users actively shared their location with their contacts. The results show that users who are central to their network are more likely to reveal most about their whereabouts. Furthermore, we show that the friend most likely to know the whereabouts of a specific individual is the one with most common contacts and/or greatest number of contacts.

**Author Keywords**
Location sharing, privacy, privacy attacks.

**ACM Classification Keywords**
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

**General Terms**
Experimentation, Human Factors.

## INTRODUCTION

The study tries to answer two important questions relating to targeted location-sharing privacy attacks. First, given a group of users and the social ties amongst them, is it possible to predict which of these users is likely to reveal the most about their whereabouts? Second, given an individual user within a particular social network, is it possible to predict which of her friends knows most about her whereabouts? To answer these, the paper presents a longitudinal study of real-time location sharing whereby the patterns of information exchange and privacy policies of a large group of users are analysed and modelled.

Real-time location sharing applications are gaining wide adoption, with a number of commercial systems now available on the market, including Foursquare, Facebook Places, and Google Latitude. Such services are frequently used in the context of online social networks (OSN), whereby one's real-time location becomes yet another sharable aspect of one's online profile. With the increasing adoption of online location sharing services, understanding the privacy implications and potential targeted attacks enabled by this new technology, becomes crucial.

A conventional approach for engineering a privacy attack is to attempt to gain ongoing access to the target's whereabouts, thereby building up a profile of that user's behaviour. In this paper, we assume that location sharing practices are likely to follow the trend of other OSN profile properties and propagate through the network of friends. The key assumption, therefore, is that a target's location can be visible to friends of friends. From the attacker's perspective, this has the benefit that they do not get "too close" to the target while still they are able to collect information about the target's location on an ongoing basis.

The two questions that this paper addresses are key in instrumenting a targeted attack against users. Such an attack would first identify a suitable target amongst a set of users. Once this has been achieved, the attacker then identifies a "weak link" in the target's list of friends. The "weak link" is a friend of the target whom the attacker will attempt to befriend in order to gain direct access to the target's whereabouts by becoming a friend of a friend. Therefore, the attacker is likely to seek for weak links who are most likely to have full access to the whereabouts of the target.


*UbiComp'11,* September 17–21, 2011, Beijing, China.

## RELATED WORK

### Sharing perceptions and strategies

Substantial research shows that people approach for developing rules and ultimately policies for sharing information with others are strongly related to the presentation of self [13], and also relate to the formulation of dialectic and dynamic behavioural mechanisms depending on circumstantial context [2] or the conjugation of disclosure, identity and temporal boundaries [24]. What was once achieved with walls, doors and other physical or architectural constraints is still to be adapted to today's communication means [34]. Privacy management is an intricate process and is further augmented in a computer mediated environment. On social networking sites, privacy regulation is a socio-technical activity involving interaction with the technological system and the group context. Individuals' privacy behaviour in such systems involves a mixture of technical and mental strategies. For instance, a technical strategy may involve the use of privacy settings to regulate content distribution to select audiences [30], while research has also shown that considering tie strength is another strategy for developing rules for disclosure [36].

Despite the evidence suggesting that users adopt objective strategies for controlling their privacy online, previous work has identified a discrepancy between people's privacy attitudes towards sharing information and their actual sharing patters [1,23]. This behaviour has been termed the "privacy paradox". For instance, a study revealed a high discrepancy between stated concerns and actual behaviour towards sharing static profile information on Facebook [1]. Other studies have further established the privacy paradox on social networking sites [34].

While it is not clear whether the privacy paradox also applies to people's perceptions towards location sharing, it certainly highlights the needs for collecting quantitative data on people's location-sharing behaviour rather than relying purely on subjective data.

### Location-sharing privacy

There is an increasing amount of work on understanding users' location-privacy needs in ubiquitous and location-aware systems relying on techniques such as diary studies [4], interviews [14], surveys [17], scenarios [16, 35] and lab and field observations [5]. Research suggests that users may start with relatively coarse and conservative preferences [28]. Over time, they take advantage of controls exposed to them and exhibit more sophisticated sharing behaviours, controlling the availability of the data through mechanisms such as disabling the service [4] or obtaining feedback about which users can see or have seen their information [14,16,28,33]. Users are also sceptical about the usefulness of location sharing in day-to-day activities, suggesting that current practices (such as calling somebody up) are sufficient [4]. However, the usefulness of such services was acknowledged in more stressful situations involving unfamiliar environments or in crisis and safety scenarios in general [14]. In such situations, information usefulness outweighs privacy concerns. Furthermore, prior work has shown that people's presence in different physical environments is likely to affect their willingness to trust and actually engage in interaction with location-based services [18].

Research investigating sophisticated privacy mechanisms, such as customizable privacy policies, has indicated that, without new interface technologies, they can present significant challenges for users. One recent study reports participants failing to implement their desired policies with a high degree of accuracy [28]. Furthermore, it also noted that although participants varied considerably in the time they spent defining their policies (between 5 and 8 minutes), the duration of this period was not strongly correlated to final policy accuracy.

It has also been observed that the recipients of the location data are typically more significant to users than the locations being shared. Perhaps unsurprisingly, users are more willing to share information with friends than acquaintances or strangers [5,33]. While recipient identity seems to be the strongest factor influencing one's willingness to share her location [10,20] time and location restrictions have been shown to also be important in capturing people's preferences [5]. Research has also shown that users are sensitive to the reactions of recipients if location information is denied or not made available [14, 28], suggesting that systems need to incorporate an element of plausible deniability. However, users do make distinctions in sharing particular locations: additional privacy is required at home when compared to work [31].

### Privacy attacks on OSNs

Targeted privacy attacks on OSNs have been demonstrated in the past. Attempts to construct social graphs for individuals from available public listings are already shown to be feasible [6]. Once achieved, social graphs can be clustered for segregating groups into sub-groups in terms of different spheres of activity for an individual [37,15]. Further results show that even hidden communities can be detected with reasonable effort [22]. This work shows that given an individual of interest, it is possible to identify a close group around that person, which may potentially be used in order to get closer to the target. To some extent, this is already done by authorities targeting criminals coordinating their activities using OSNs [7,9], and it usually involves some level of active probing [29] which in the context of OSNs may mean striking friendships with individuals close to the target so as to avoid detection.

The characteristics of privacy attacks in the context of location sharing differ from privacy attacks online social networks because of two reasons. Location-sharing applications include information about users' physical

whereabouts, which can lead to access to one's physical self. Empirical evidence show that users fear that revealing their location to people they do not trust may lead to physical and property harm [32]. Furthermore, users' decisions on location sharing may differ considerably than decisions taken in the context of social networks, making this subject worthwhile of investigation.

#### Identifying "weak links"

Prior work on social networks may be used to derive some hypotheses about who is likely to share information with whom on a social network. For instance, Petronio's theory of Communications Privacy Management (CPM) describes an iterative process of rule development, boundary coordination and boundary turbulence [25]. Rule development can be defined as the process of developing regulations about who to tell what. These regulations guide our everyday disclosures, and are a function of our context and disclosure goals. Ties of differing strength have varying disclosure norms, thus Stutzman theorizes rule development is a function of network composition [30]. For example, a network that is more heavily focused on strong ties may require higher levels of privacy, as disclosures among strong ties are more personal in nature [36]. This suggests that network structure may be used as a basis for attempting to predict disclosures amongst individuals.

Recent work on sharing ephemeral information shows that rule development is a function of tie strength [27]. In order to test CPM's rule development process on the context of posting content to Facebook, users were presented various scenarios of information disclosure and were prompted to decide how and with whom to share that information with. Results show users are more prone to share with stronger ties as opposed to weak ties. These findings where uniform across the various scenarios of information disclosure presented to participants. Intended and expected audience for both profile and ephemeral information was a function of tie strength [27,30]. Both authors report that users' perceived audience for the information they share is mostly composed of strong ties.

## STUDY

#### Definitions

The following are definitions of metrics used in the study that follows.

- *Social Graph*: A set of individuals and the explicit friendship ties amongst them.
- *Degree Centrality* : The degree centrality of a user is the number of direct connections (or "friends") that the user has in the social graph. These were the friends of the user on Facebook that were also users of Locaccino.
- *Betweenness Centrality* : The betweenness centrality of a user is the number of shortest paths between all pairs of nodes in the social graph that pass through the node representing the user. For a more thorough description of the betweenness centrality, the reader is directed to [11].
- *Openness* : The *openness* of the ordered pair (A, B) of users is the percentage of simulated location requests made to A by B that were granted by A's policies.
- *Trust* : The *trust* of a user A is the mean of the openness values (A, B) where B ranges over all of A's friend. i.e. it is the average openness of user A towards all her friends.
- *Trustworthiness* : The trustworthiness of a user A is the mean of the openness values (B, A) where B ranges over all of A's friends. i.e. it is the average openness of A's friends towards A.
- *Trust Rank :* Given a user A and a user B who is a friend of A, the *trust rank of B with respect to A* is *i* if there are precisely *i-1* friends $C_1, C_2 \ldots C_{i-1}$ of A such that the openness of $(A, C_j)$ , $1 \leq j < i$, is greater than the openness of (A, B). i.e. the trust rank is obtained by ranking A's friends in terms of how much they are trusted by A.
- *Degree Rank* : Given a user A and a user B who is a friend of A, the *degree rank of B with respect to A* is *i* if there are precisely *i-1* friends $C_1, C_2 \ldots C_{i-1}$ of A such that the degree centralities of $C_1, C_2 \ldots C_{i-1}$ are greater than that of B. i.e. the degree rank is obtained by ranking A's friends in terms of their degree centralities.
- *Mutual Rank* : Given a user A and a user B who is a friend of A, the *mutual rank of B with respect to A* is *i* if there are precisely *i-1* friends $C_1, C_2 \ldots C_{i-1}$ of A such that the number of common friends A has with each of $C_1, C_2 \ldots C_{i-1}$ is greater than the number of common friends that A has with B. i.e. the mutual rank is obtained by ranking A's friends in terms of how many mutual friends they have with A.

#### Hypotheses

Previous work suggests a relationship between social network structure, tie strength and the patterns of disclosure amongst individuals (e.g. [30]). In attempting to identify which individual is more likely to reveal information about their whereabouts, one may hypothesise that individuals who are more central to the network are more likely to do so. A possible explanation would be that such individuals are more likely to engage in collaboration and coordination activities, therefore it may be more likely that they are willing to share their real-time location with others. This reasoning provides ground for the first experimental hypothesis:

- H1: Individuals who are more central to the social graph are likely to reveal the most about their location.

Upon determining a suitable person to target, the next step in a potential attack would be to befriend someone from the target's social network. Considering that previous literature

suggests that reciprocity is an important driving force in social networks [26], one can expect that the target is likely to share their location with someone in the social network out of their desire to reciprocate. Hence, the friend of the target with the most number friends, who by means of H1 is likely to share their own location, is someone with whom the target may wish to share their location in order to reciprocate. That person is therefore a potential "weak link" whom the attacker might befriend in order to get closer to the target. This leads to the second experimental hypothesis:

- H2: The target's friend with the highest degree has higher probability of knowing more about the target.

Finally, it can be argued that shared membership and being part of the same community would be suggestive of two individuals who may be possibly involved in joint activities requiring coordination. In addition, literature on homophily has shown that individuals who share mutual friends are more likely to be alike, thus likely to engage in joint activities [21,8]. It is therefore plausible to hypothesise that individuals who belong to the same group are more likely to share their real-time location with each other, thus becoming candidate "weak links". This leads to the third experimental hypothesis:

- H3: The target's friend with most common ties with the target knows most about the target.

To test these hypotheses, the location-sharing system described next was deployed and used by a large group of users longitudinally.

### System

The study was conducted by deploying Locaccino, a real-time location sharing application integrated in the OSN Facebook. The application consists of two components: a Web application component and a mobile component. Various version of the mobile component were developed to run on multiple mobile platforms: windows and apple laptops, and Symbian smartphones. The purpose of this component is to collect in real-time a user's location and then upload it to a central server.

The Web application component of Locaccino (Figure 1) allows users to set preferences regarding how their location is shared with their Facebook contacts. Users are given the option to create policies in order to manage their location sharing. Policies specify the conditions under which the location should be revealed to another user. These conditions include the identity of the recipient of the information, the time and day, and the actual location where the user is. For instance, one may specify a policy to allow work colleagues to obtain one's location only during work hours and when they are in town.

Participants were recruited on campus using advertisements on-line and via email, as well as through national press covering the features of our system.

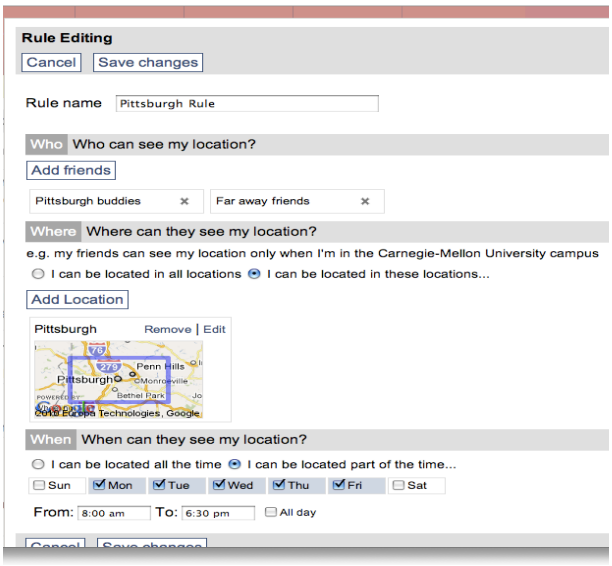


**Figure 1. Screenshot of Locaccino's functionality that allows users to construct their location sharing policy rules.**

The system was used longitudinally and more than 300 users installed the application and actively begun using it to share their location with colleagues. For the purposes of the study presented here, the following information was collected about users:

- Social graph: An undirected unweighted graph describing the friendship between all the participants. In this graph, a node represents a user, and two nodes are connected if they are friends on Facebook.
- Policy graph: A directed weighted graph describing the privacy policies between the users. In this graph each node represents a user, and user A is connected to user B if user if these two users are friends on Facebook. In addition, the weight of the edge from user A to user B is a value between 0 and 1 based on the "openness" of user A towards user B. The weight of the opposite edge, i.e. the openness of user B towards user A, is independent and may be a different.

The openness value of (A,B) was calculated as the percentage of B's possible requests that were granted by A's policies. The openness value from one user towards another represents the extent to which a user is willing to share their location with another user. In our case we rely on users' policies to capture and quantify this feature. Specifically, to generate a value representative of the openness between two users we conducted the following procedure. For each pair of users (A,B) in the dataset we ran a simulation whereby user B repeatedly requested the location of user A. These simulated requested were processed by the policies of user A, and the result was either positive or negative, thereby either showing or hiding user A's location respectively. During this analysis the

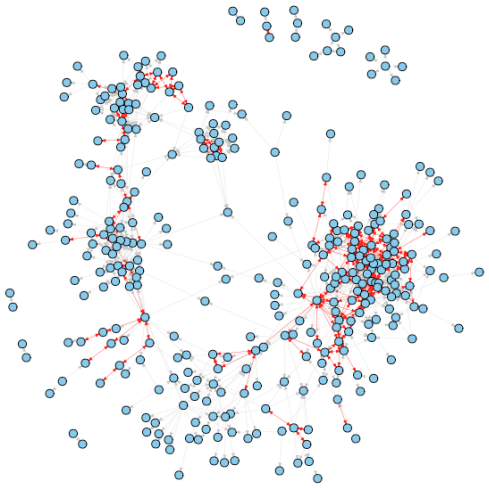

**Figure 2. The graph representing the participants (nodes) and their trust relationships as directed edges. Mutually open relationships are highlighted in red.**

movement of user A was the same as recorded during the study.

## RESULTS

The study ran for a month with 340 users who were already users of Facebook. The derived policy graph contained 1778 policy rules, two for each of the 889 friendship ties within the user population (Figure 2).

Each policy described the openness of one user towards other users, ranging from 0 to 1. For each user the average openness that they show towards their friends was calculated (referred to as "trust" towards others) and is summarised in Figure 3, while the average openness that a user was shown by his friends (i.e. their "trustworthiness") is shown in Figure 4.

### Hypothesis testing

*H1: Individuals who are more central to the social graph are likely to reveal the most about their location.*

A Kruskal_Wallis non-parametric test of independent samples [e.g. 12] showed that there was a significant effect of a node's betweenness on that node's trust towards its direct connections ($H(45)=82.111$, $p<0.001$) but not on that node's trustworthiness ($H(45)=56.168$, $p=0.123$). Furthermore, there was a significant effect of degree centrality on node trust ($H(23)=82.076$, $p<0.0001$) and also on node trustworthiness ($H(23)=35.276$, $p<0.05$).

*H2: The target's friend with the highest degree has higher probability of knowing more about the target.*

To test this hypothesis all users with less than 2 friends in the dataset were discarded from the analysis, leaving 247 users. This data was discarded because no comparison can

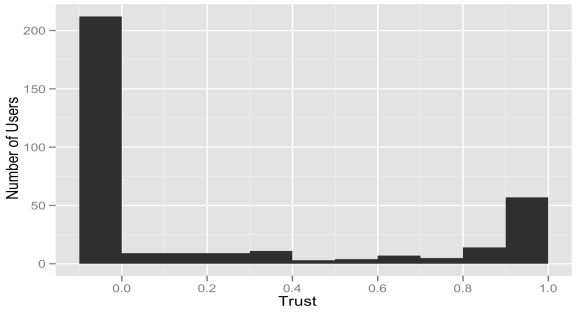


**Figure 3. Histogram of distribution of nodes' average openness (i.e. the average of all outgoing ties for each node)**

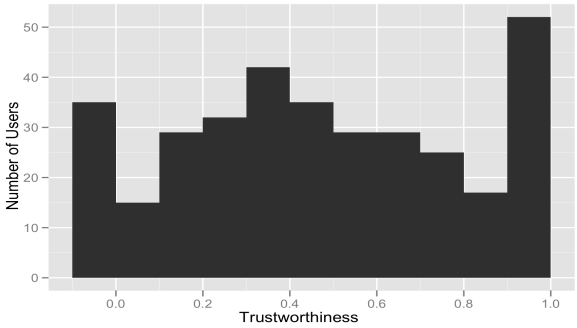


**Figure 4. Histogram of nodes' average trustworthiness (i.e. the average of all incoming ties for each node).**

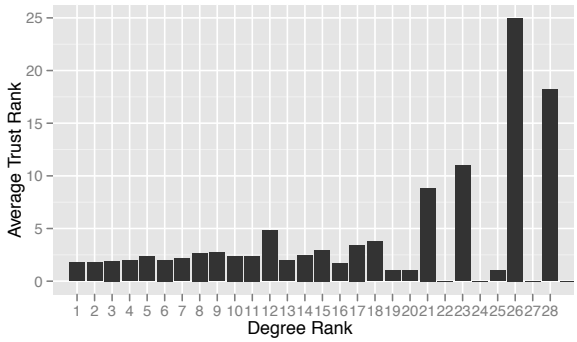


**Figure 5: Degree rank of nodes (x-axis) versus the average trust rank (y-axis) for all nodes of a specific degree rank**

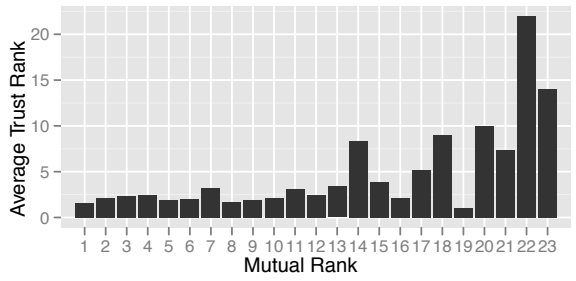


**Figure 6. Histogram of Mutual rank (x-axis) vs. average trust rank (y-axis) for all nodes of a specific CommonFriends rank.**

be carried out for users with a single friend. For each user A, all of A's friends were ranked in terms of how much they are trusted by A (Trust Rank), and in terms of how many friends they actually have (Degree Rank). This gave for each friendship relationship in the data two values: Trust Rank and Degree Rank respectively (Figure 5). A chi-square test showed a significant relationship between Degree Rank and Trust Rank ($x^2=3981.723$, $dF=744$, $p<0.001$) while there was a positive correlation between the two variables (0.239, $p<0.01$).

*H3: The target's friend with most common ties with the target knows most about the target.*

To test this hypothesis all users with less than 2 friends were discarded from the analysis, leaving 247 users. For each user A, all of A's friends were ranked in terms of how much they know about A (Trust Rank), and in terms of how many mutual friends they have with A (Mutual Rank). This gave us for each friendship relationship in the data two values: Trust Rank and Mutual Rank respectively (Figure 6). A chi-square test showed a significant relationship between Mutual Rank and Trust Rank ($x^2=3210.841$, $dF=682$, $p<0.001$), while there was a positive correlation between the two variables (0.252, $p<0.01$).

**Structural analysis**

Finally, an analysis was conducted to assess the extent to which friends with the highest degree are the same as friends with a large number of common friends. A chi-square test showed a significant relationship between Degree Rank and Mutual Rank ($x^2=6548.051$, $dF=528$, $p<0.001$), and a positive correlation between Degree Rank and Mutual Rank of 0.81 ($p<0.01$) as shown in Figure 7.

Furthermore, a triad analysis was conducted, to assess the extent to which there exists a bias in how trust and trustworthiness was distributed across the network. The analysis was conducted by first classifying each bi-directional edge in one of three possible states: balanced-high (meaning both people are sharing in full or partially), balanced-low (meaning that both people are not sharing), and unbalanced (meaning that one person is sharing while the other is not). Given the three possible labels for each edge, there exist 10 possible "templates" for triads, depending on the combination of its bidirectional edges (see Table 1). Each triad in the graph was labelled appropriately, and the frequency of occurrence of each template was calculated.

In addition, for each of the 10 templates the theoretical expected frequency of occurrence was calculated, as described in [19], by assuming that the same edges were randomly distributed on a graph with identical topography. The relationship between the observed and expected frequency for each of the 10 templates is shown in Figure 8. The figure shows a modest correlation ($R^2=0.75$), with the exception of the data point at (119,225) corresponding to

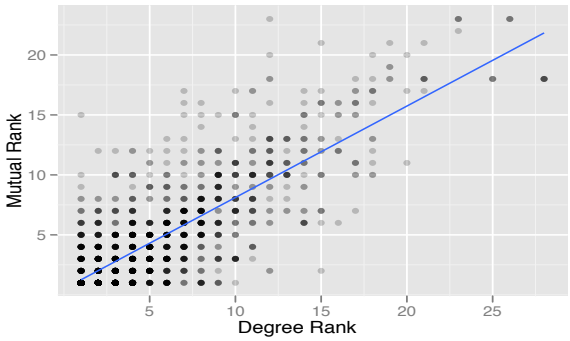


**Figure 7. X-axis: degree rank of a neighbour wrt another node. Y-axis: common friends rank of a neighbour wrt another node. Darker dots indicate overlapping points.**

| Template | Expected Frequency | Observed frequency |
|---|---|---|
| 1 | 292 | 209 |
| 2 | 142 | 119 |
| 3 | 119 | 225 |
| 4 | 23 | 22 |
| 5 | 39 | 10 |
| 6 | 1 | 19 |
| 7 | 16 | 5 |
| 8 | 3 | 1 |
| 9 | 3 | 0 |
| 10 | 0.7 | 28 |

Unbalanced Balanced-Low Balanced-High

**Table 1. Expected frequency (given a random model) and observed frequencies for each of the 10 possible triad templates.**

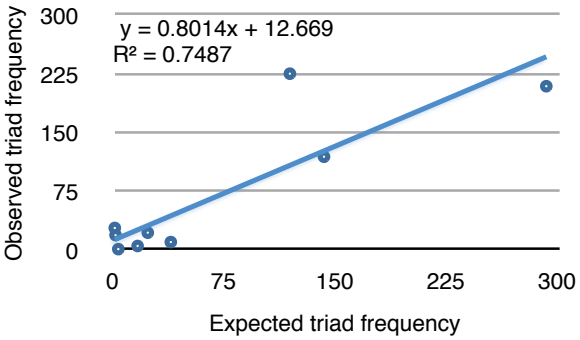


**Figure 8. Correlation between the expected and observed frequencies for each of the 10 possible triad templates.**

the template "Balanced-high, Unbalanced, Unbalanced". Removal of this point would substantially improve the correlation (R^2=0.96).

## DISCUSSION

### Targeted location-sharing privacy attacks

This paper proposes a threat model related to location-sharing privacy, whereby the attacker attempts to collect data about a target's longitudinal movements. To do this, the attacker first needs to identify suitable targets such that his chances of success are maximised. Once a suitable target is identified, then the attacker attempts to gain access to the target in order to collect data about the target's location, but not "too close" to avoid detection. Therefore, even though a strategy for collecting data on a target's whereabouts would be to attempt to befriend the target directly, that increases the chances of the attacker being noticed. This paper assumes that the attacker can attempt to collect data about the target by befriending one of the target's friends, i.e. a "weak link". This will make the attacker a "friend of a friend" of the target, which is arguably adequate to gain access to the target's location. In order to achieve this, the attacker needs to figure out which of the target's friend are more likely to have access to the target's location data, and are therefore a more suitable person to befriend. The results from Figure 3 show that, on average, nodes exhibit a bimodal distribution of trust which is weighted towards not sharing at all. Hence, if the attacker picks a target's friend at random, they are about four times more likely to not gain access at all to the target's data.

To assess the extent to which such an attack can be engineered, the study presented here answers two important questions relating to this kind of targeted location-sharing privacy attacks. First, given a group of users and the friendship ties amongst them, is it possible to predict which of these users is likely to reveal the most about their whereabouts? Second, given an individual user within a particular social network, is it possible to predict which of her friends know most about her whereabouts?

The study presented here captured a measure of "openness" between individuals, which reflects the probability that a request for someone's real-time location is likely to be satisfied. An advantage of using a generic measure, which we refer to us trust (when a person of interest is open towards someone else) and trustworthiness (when someone else is open towards a person of interest), is that it can be applied across multiple features of online social networks. Therefore, while commercial location-sharing systems vary in features and their capabilities evolve over time, the measure of trust and trustworthiness is likely to remain an underlying driver in guiding users' decision to share their location with others.

### Identifying a suitable target

The motivation for H1 was to suggest a way in which the attacker can identify users who are more likely to share their location with friends. The hypothesis was that individuals who are more central to the social network reveal the most about themselves, motivated by the observation that such individuals are more likely to engage in collaboration and coordination activities. Our results suggest that a user's network centrality as measured by betweenness and degree centrality had a significant effect on the amount of trust that user was willing to show towards their friends, thus supporting hypothesis.

The results suggest that individuals who are more central to their network are more likely to be willing to share their location with others, and therefore they are good targets for a potential attacker. Hence, an attacker can conduct a basic analysis of the network structure to identify central nodes, and then attempt to target more central nodes since they are more likely to share their location. It can be argued that individuals who are more central to the network are more socially active, and maintain more social relationships. This is likely to require them to take part in more social activities, and therefore it can be argued that these conditions require more coordination on their part. This offers one explanation as to why the findings in this study suggest that more central users did in fact share their location more often.

### How to target individuals

Once the attacker has identified a target who is likely to be open and share their location, the next step is to develop a strategy for targeting that individual. The threat model discussed in this paper assumes that the attacker will not attempt to befriend the target directly, since that bears a high risk of being detected. Instead, the attacker can attempt to befriend someone from the target's friends since that can give them access to the target's location data without bringing them "too close" to the target. Therefore, the next step for the attacker is to identify a "weak link" in the target's list of friends, or a person who is likely to be granted access to the target's location data. The study tested two possible strategies for identifying weak links: based on the number of friends that a weak link may have (H2), and based on the number of common friends that the weak link may have with the target (H3).

Prior studies have shown the importance of reciprocity in social interactions, thus providing the motivation for H2. More specifically, studies have shown that when an individual performs a favour or act that bestows trust upon another individual, that individual is likely to feel obliged to reciprocate the favour or act. The motivation for H2 comes from this perceived obligation and from H1. The results show a positive correlation between the amount of trust that an individual bestows on each of his friends and the number of friends of those friends. This suggests that the attacker

can identify a suitable "weak link" of the target by considering the target's list of friends and identifying those individuals with the highest number of friends of their own. Such individuals are more likely to be social active, and are therefore more likely to choose to share their location with the target (see H1). The target, by virtue of reciprocity, is therefore more likely to share their location with such individuals.

A competing, and possibly complementary hypothesis for identifying weak links is H3, which states that the target's friend with most common ties with the target knows most about the target's whereabouts. This can be due to the fact that the existence of common friends can indicate shared membership in a community or organisation. The results show that there is a significant positive correlation between the trust of a target towards each of his friends and the rank of that friend in terms of the number of mutual friends he has with the target. The results provide a clear strategy for how an attacker can identify a "weak link", which entails identifying who from the target's list of friends has the highest number of common friends with the target. One explanation for these findings is that individuals who share many friends, and are thus likely to belong to the same social groups, are more likely to share their location in order to coordinate their activities better, as well as to maintain an increased awareness of each other's ongoing activities.

Finally, it should be pointed out that the analysis provides evidence that H2 and H3 are directly related. Since both hypotheses were supported by the analysis, this is not surprising. The results show a strong positive correlation between H2 and H3 in that a target's friends who have many friends are also likely to have a lot of common friends with the target. One explanation for this relationship may be that individuals who have many friends of their own are more likely to be extroverts who socialise and engage in multiple social interactions activities. Their behaviour therefore increases the likelihood of them being friends with mutual friends with the target simply because they have a lot of friends.

#### Triads and small group privacy

The results of the structural analysis presented here offer insights into how, in the context of location-sharing, triads of users distribute and balance trust and trustworthiness. In addition to being useful in understanding the behaviour of our participants, these results are also useful in situations where only partial information may be known about the network.

The structural analysis shows that even though under a completely random model we expected to observe only one triad where all three members trust each other (template 10 in Table 1), we actually observed 28 such triads. Furthermore, the correlation analysis highlights triad template 3 as being substantially different from the overall correlation pattern between expected and observed frequencies. In this case, this result shows that we observed quite often situations where two people trust each other but both maintain unbalanced relationships with a third individual. This is a balanced situation and expected to be more frequent in a realistic setting than in a purely random environment [e.g. 11].

The results from this analysis coincide with prior work in that people tend to avoid unbalanced situations and prefer the comfort of balanced triads . Furthermore, these results can be used to make predictions in situations where incomplete information has been collected about individuals. This is possible since given three individuals and 2 of the 3 relationships between them, we may be able to predict the third relationships. For example, given a triad with two balanced-high relationships, the chances of the third relationship being balanced-low is very close to zero, unbalanced is 15%, and balanced-high is 85%.

#### Protection against such privacy attacks

The attack described here assumes that the attacker is trying to gain longitudinal access to the target's whereabouts, and does so by avoiding detection since they do not need to befriend the target directly, but only one of their friends. Assuming that on average users have about 150 friends in a social network, then the attacker's strategy ensures that he is one of about 22000 people who are friends-of-friends of the target, making detection much harder.

One strategy that the platform could follow in case of a pull-based location-sharing model would be to ensure that individuals are notified if anyone is making too many location-sharing requests. This could be implemented in the form of a user-defined threshold or as a nudging mechanism intended to help people refine their sharing preferences [3]. In the case of a push-based model, the users can ensure that their information is visible only to their friends directly, and to no-one beyond that. Similarly, limits could be imposed on how often a user can update their location, hence offering an upper bound on how much users can reveal about their whereabouts. However, such solutions seem to contradict the needs of commercial systems which appear to strive for increasing the amount of shared information.

#### Making useful predictions

While the work described here was framed in the context of a privacy attack, the hypotheses that were tested may be useful in developing user-friendly features that can automatically provide useful suggestions to users. For instance, the hypotheses discussed earlier provide an indication on how to identify a person who is likely to know the whereabouts of an individual of interest. It may be the case that the individual of interest has not logged into the location-sharing system to update their location, due to technical difficulties, time constraints, or any other plausible difficulty. Under such circumstances, the system

may be able to make automated suggestions about who to ask regarding the whereabouts of the person of interest based on a simplistic network-structure analysis. Therefore, in cases of high urgency it is possible to offer such recommendations as a fall-back strategy.

### Limitations of the study

In a realistic environment there may be multiple factors affecting the sharing of information, many of which are inadvertently manipulated by users. For instance, battery life and group norms may be important factors that urge users to hide or share their location. These were not taken into account in this study.

Furthermore, this study presents and tests a generic strategy for engineering such an attack. Clearly, the fine details of the social platform where this information is recorded and shared are important, and may facilitate or hinder the success of such an attack. For instance, being friend of a friend may be "too close" or "too far" to obtain location information, while some auditing mechanism may allow users to see who is viewing their location information repeatedly.

Finally, it is important to take into consideration here the fact that users of this location sharing application start with a default privacy policy of not sharing their location information with anybody in the network. We cannot rule this out a contributing factor to our result that more central nodes trust more, as they are also likely to be seasoned users of the system and hence have invested more time to articulate their location sharing preferences.

## CONCLUSION

This paper presents a study that aims to answer two important questions related to targeted location-sharing privacy attacks: (1) given a group of users and their social graph, is it possible to predict which among them is likely to reveal most about their whereabouts, and (2) given a user, is it possible to predict which among her friends knows most about her whereabouts.

The results show that users who are more central to their social network (both locally and globally) are more likely to share information about their location, and hence are more "vocal". In addition, the findings show that given a target, that target's friend who either has many friends or many common friends with the target is more likely to be trusted by the target.

The findings of this study are important both in understanding how privacy attacks can be engineered and how they can be prevented. An important next step for this work is the application of these insights for the development of automated protection and suggestion mechanisms that will make the sharing of real-time location safer and more useful.

## ACKNOWLEDGEMENTS

This work is funded by NSF grants CNS-0627513, CNS-0905562, CNS-1012763, by CyLab at Carnegie Mellon under grants DAAD19-02-1-0389 and W911NF-09-1-0273 from the Army Research Office. Additional support has been provided by Google and the CMU/Portugal Information and Communication Technologies Institute and the Portuguese Foundation for Science and Technology (FCT) grant CMU-PT/SE/0028/2008 (Web Security and Privacy). The authors would also like to acknowledge the entire Locaccino team, including Jason Hong, Lorrie Cranor, Paul Hankes Drielsma, Justin Cranshaw, Patrick Gage Kelley, Jialiu Lin and Michael Benisch for their contributions.

## REFERENCES

1. Acquisti, A. and Gross, R. (2006). Imagined communities: awareness, information sharing, and privacy on the Facebook. Proc. PET 2006, Springer, 36-56.
2. Altman, I. (1977). Privacy Regulation: Culturally Universal or Culturally Specific? Journal of Social Issues, 33 (3), 66-84.
3. Balebako, R., Leon, P.G., Mugan, J., Acquisti, A., Cranor, L.F., Sadeh, N. (2011) Nudging Users Towards Privacy on Mobile Devices, CHI 2011 workshop article, May 2011
4. Barkhuus, L. (2004). Privacy in location-based services, concern vs. coolness. Mobile HCI 2004 workshop: Location System Privacy and Control.
5. Benisch, M., Kelley, P.G., Sadeh, N., Cranor, L.F., (2010). Capturing Location-Privacy Preferences: Quantifying Accuracy and User-Burden Tradeoffs. Personal and Ubiquitous Computing (PUC). Forthcoming.
6. Bonneau, J., Anderson, J., Anderson, R. and Stajano, F. (2009) Eight Friends Are Enough: Social Graph Approximation via Public Listings. SNS'09.
7. Choo, K-K. R. and Smith, R. G. (2008). Criminal Exploitation of Online Systems by Organised Crime Groups. Asian Criminology (2008) 3:37–59
8. Christakis, N., Fowler, J.H. (2008). The collective dynamics of smoking in a large social network. The New England journal of medicine 358(21): 2249-58.
9. CISC (2010) Report on Organized Crime. Criminal Intelligence Service Canada. Available At: http://www.cisc.gc.ca [Last Accessed 25th October 2010]. ISBN 978-1-100-51931-9.
10. Consolvo, S., Smith, I. E., Matthews, T., LaMarca, A., Tabert, J., and Powledge, P. (2005). Location disclosure to social relations: why, when, & what people want to share. CHI 2005, 81-90.

11. Easley, D., and Kleinberg, J. (2010). Networks, Crowds, and Markets: Reasoning About a Highly Connected World. Cambridge University Press.
12. Gibbons, J.D. (1993). Nonparametric statistics: An introduction. Sage University Paper series on Quantitative Applications in the Social Sciences, 07-090.
13. Goffman, E. (1959). The presentation of self in everyday life. Garden City, NY: Doubleday Anchor.
14. Hong, J. I. and Landay, J. A. (2004) An architecture for privacy-sensitive ubiquitous computing. MobiSys '04, 177-189.
15. Jones, S., and O'Neill, E. (2010). Feasibility of structural network clustering for group-based privacy control in social networks. In Proceedings of the Sixth Symposium on Usable Privacy and Security (SOUPS '10). ACM, New York, NY, USA, , Article 9 , 13 pages
16. Kelley, P. G., Hankes Drielsma, P., Sadeh, N., and Cranor, L. F. (2008). User-controllable learning of security and privacy policies. AISec 2008, 11-18.
17. Khalil, A. and Connelly, K. (2006). Context-aware telephony: privacy preferences and sharing patterns. CSCW '06, 469-478.
18. Kostakos, V. and Oakley, I. (2009). Designing Trustworthy Situated Services: an Implicit and Explicit Assessment of Locative Images' Effect on Trust. CHI, Boston, USA, pp. 329-332.
19. Kostakos, V. and Venkatanathan, J. (2010). Making friends in life and online: Equivalence, micro-correlation and value in spatial and transpatial social networks. IEEE SocialCom, Minneapolis, USA, pp. 587-594.
20. Lederer, S., Mankoff, J., and Dey, A. K. (2003). Who wants to know what when? privacy preference determinants in ubiquitous computing. CHI 2003, ACM Press, 724-725.
21. McPherson, Miller, Lynn Smith-Lovin, and James M Cook. "Birds of a Feather: Homophily in Social Networks." Annual Review of Sociology 27(1):415-444.
22. Nagaraja, S. (2008) The economics of covert community detection and hiding. WEIS: Workshop on the Economics of Information Security.
23. Norberg, Patricia A, Daniel R. Horne, and David A. Horne (2007). The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors. Journal of Consumer Affairs, 41(1):100-126.
24. Palen, L., Dourish, P. (2003). Unpacking Privacy for a Networked World. In Proc. of the Conference on Human Factors and Computing Systems: CHI 2003, ACM Press 129–136.
25. Petronio, S. (2002). Boundaries of Privacy: Dialectics of Disclosure. SUNY, Albany, NY
26. Regan, D.T. (1971). Effects of a favor and liking on compliance. Journal of Experimental Social Psychology, 7(6):627-639.
27. Reynolds, B., Venkatanathan, J., Goncalves, J., and Kostakos, V. (2011). Sharing Ephemeral Information in Online Social Networks: privacy perceptions and behaviours. In proceedings of INTERACT.
28. Sadeh, N., Hong, J., Cranor, L., Fette, I., Kelley, P., Prabaker, M., and Rao, J. (2008). Understanding and capturing people's privacy policies in a mobile social networking application. Personal and Ubiquitous Computing 13, 6, 401-412.
29. Shaikh, S. A., Chivers, H., Nobles, P., Clark, J. A. and Chen, H. (2008). Network reconnaissance. Network Security, 2008(11):12-16.
30. Stutzman, F., Kramer-Duffield, J. (2010) Friends only: examining a privacy-enhancing behavior in Facebook In Proc. of the Conference on Human Factors and Computing Systems: CHI 2010, ACM Press 1553-1562.
31. Toch, E., Cranshaw, J., Drielsma, P.H., Tsai, J. Y., Kelley, P. G., Cranor, L., Hong, J., Sadeh, N. (2010) Empirical Models of Privacy in Location Sharing, in Proceedings of the Twelfth International Conference on Ubiquitous Computing. Ubicomp 2010
32. Tsai, J., Kelley, P.G., Cranor, L.F., and Sadeh N. (2010). Location- Sharing Technologies: Privacy Risks and Controls. Journal of Law and Policy for the Information Society, 2010.
33. Tsai, J.Y., Kelley, P., Drielsma, P., Cranor, L.F., Hong, J., and Sadeh, N. (2009). Who's viewed you?: the impact of feedback in a mobile location-sharing application. CHI '09, 2003–2012.
34. Tufekci Z. (2008). Can You See Me Now? Audience and Disclosure Management in Online Social Network Sites. Bulletin of Science and Technology Studies. Volume 11, Number 4, June 2008 , pp. 544-564(21).
35. Wagner, D., Lopez, M., Doria, A., Pavlyshak, I., Kostakos, V., Oakley, I., Spiliotopoulos, T. (2010). Hide And Seek: Location Sharing Practices With Social Media. MobileHCI '10, 55-58.
36. Wellman, B. and Wortley, S. (1990). Different Strokes from Different Folks: Community Ties and Social Support. American Journal of Sociology 96, 3, 558-588.
37. Xu, X., Yuruk, N., Feng, Z. and Schwieger, T. A. J. (2007) SCAN: a structural clustering algorithm for networks. Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining, 824-833.

## 4.3 Critical review of the work and conclusion

***Revisiting the Kruskal-Wallis Test for Hypothesis H1***

In the results section of the paper we reported Kruskal-Wallis tests in order to examine the validity of hypothesis H1. Here we revisit that analysis in order to elaborate on the procedure and to address an unclarity in the analysis.

The hypothesis we wish to test, hypothesis H1, is the following : "*Individuals who are more central to the social graph are likely to reveal the most about their location"*. For measures of centrality, we use betweenness centrality and degree centrality. To measure how much each user reveals about their whereabouts to their friends, we use the "trust" metric. These metrics are defined in the "Study" section of the paper.

Since the distribution of trust is heavily bimodal (Figure 3 of the paper), we cannot use a correlation analysis to examine whether trust is related to centrality. Therefore, we code trust into a categorical variable with 3 levels : Low trust (trust < 0.2), Moderate trust (0.2 >= trust <= 0.8) and High trust (trust > 0.8). An alternate coding scheme we can use is : No trust (trust=0), Partial trust, and Full trust (trust=1). Both coding schemes lead to the same outcomes. We present the analysis based on the Low, Moderate, High coding scheme here.

The analysis reported in the Results section of the paper treats the centrality variable as categorical variable. For example, degree centrality is treated as a categorical variable with 24 levels because there are 24 unique degree centrality values that the nodes in the network have. Similarly, betweenness centrality (converted to log scale and rounded off to the nearest integer) takes 46 unique values and is therefore treated as a categorical variable with 46 levels. Thus the Kruskal-Wallis tests reported in the paper with centrality as the categorical variable tests whether the trust values significantly differ for different levels of centrality scores. Due to the large number of levels of centrality scores (24 for degree and 46 for betweenness), this analysis can be difficult to interpret in terms of how the trust scores vary with centrality. Therefore, here we provide additional results that examine the association of trust with centrality by reporting the Kruskal-Wallis test with trust as a categorical variable and centrality as a continuous variable. This effectively checks whether centrality scores significantly differ between

the three levels of trust based on our coding scheme. If higher levels of trust are associated with higher centrality, we take that as evidence in support of H1.

First, the test shows a significant effect of trust level on degree centrality (Kruskal-Wallis $\chi^2 = 51.14$, df = 2, $p<0.001$). The average degree centrality of the nodes within the network for the three levels of trust are as follows - Low Trust : 7.8, Moderate Trust : 15, High Trust : 16.5. In addition, there is a significant effect of trust level on betweenness centrality (Kruskal-Wallis $\chi^2 = 51.22$, df = 2, $p<0.001$). The average betweenness centrality for the three levels of trust are as follows - Low Trust : 612.7, Moderate Trust : 1847.5, High Trust : 2198.7 .

Another metric we mention in the paper is "trustworthiness". This metric captures for each participant how much her friends revealed their location to her on average. The trustworthiness metric is likely to be an unreliable metric for understanding the relationship between participants' centralities and how likely others are to reveal their location to them. This is due to the fact that if more central nodes are also more trusting in general, as suggested by our analysis on trust, then nodes that have low centrality but are connected only to one or a few central nodes might be likely to have high trustworthiness scores. Similarly, a central node might have a low trustworthiness score in comparison with the other nodes across the entire network if her ties are not trusting in general, but this does not imply that her ties trust her less than they trust their own other ties. In order to understand the relationship between participants' centrality and how likely they are to be trusted, we need to understand, for example, how likely a given node is to trust a tie with high centrality in comparison with ties of lower centralities. This is broadly the method of analysis we have adopted in the subsequent parts of the Results section to test hypotheses H2 and H3, rather than rely on the aggregate measure of trustworthiness.

***Methodological Potential***

In this work we attempted to examine disclosure behaviour and online social network structure in conjunction. By using the privacy policy (or privacy settings) of users in the online social networking system, we were able to generate an "openness" metric, that reflects disclosure behaviour for users across the system. This simple idea enables us to study disclosure at the much larger scales at which we are typically able to study network structure from the graphs of online social networks today. Moreover, this

disclosure information can be embedded into the network structure as link attributes, thus enabling us to use techniques from social networks analysis to draw rich insights into the behaviour patterns of users.

Our analysis of the Locaccino sample is only one case in the application of these ideas, where we attempted to demonstrate how disclosure information and social network graphs can be analyzed together. Thus, this work can be considered to be attempt towards establishing a general method that can be replicated in other systems. While we analyzed the disclosure patterns of a few hundred users, the same techniques can be used to study samples of thousands and even millions of users of a social networking system. The first step in such an analysis requires that we are able to generate an "openness" metric as we have done in this study, either from the privacy settings or from the history of sharing behaviour of users. From here, we can compute metrics such as trust, trust rank, mutual rank and so on to test various hypotheses we might formulate based on the social networking system. We can also examine the dynamics of micro-level groups with respect to disclosure behaviour, such as the triad analysis we carried out for our sample. Thus, based on our experience from this study, we suggest that there is much potential for large scale combined analyses of disclosure patterns and social network structure within online social networking systems.

#### *Conclusion*

This work brings together disclosure behaviour and online social network structure by examining the relationship between individuals' positions in the network and the disclosure of location information in a location sharing social network. The results provide insights on which users within the social network structure are likely to disclose most about their whereabouts and which users are likely to know most about others' whereabouts in the social network. Further structural analyses shed light on the dynamics of small groups with respect to location disclosure behaviour. The study demonstrates the potential of our methodology for the combined analyses of disclosure behaviour and social network structure on a large scale across different online social networking systems.

## Chapter References

The following are the references cited in this chapter outside of the embedded conference paper.

Aral, S., Muchnik, L., Sundararajan, A. (2009). Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. Proc Natl Acad Sci USA 106(51):21544–21549.

Bazarova, N. N. (2012). Public intimacy: disclosure interpretation and social judgments on Facebook. J. Commun., 62, 815–832.

Leskovec, J., Huttenlocher, D., & Kleinberg, J. (2010). Predicting positive and negative links in online social networks. In Proceedings of the 19th international conference on World wide web (pp. 641-650). ACM.

Stutzman, F. and Kramer-Duffield, J. (2010) Friends Only: Examining a Privacy-Enhancing Behavior in Facebook. Proc. Conf. Human Factors and Computing Systems: CHI 2010, pp. 1553–1562. ACM Press.

Toch, E., Cranshaw, J., Hankes-Drielsma, P., Springfield, J., Kelley, P., Cranor, L., Hong, J. and Sadeh, N. (2010). Locaccino: A privacy-centric location sharing application. In Proceedings of the 12th ACM international conference adjunct papers on Ubiquitous computing-Adjunct (pp. 381-382). ACM.

# 5. Network Structure, Personality and Social Capital

## 5.1 Introduction

This chapter presents paper (c) "*Network, Personality and Social Capital"* published in the proceedings of the ACM conference on Web Science, 2012. This is the first of two studies looking into the relationship between online network structure and social capital. Our motivation for the work in this chapter was to touch upon two related issues regarding the benefits one receives from his or her social ties. The first is a body of work in traditional social science which suggests that the structure of the social ties around individuals affects the benefits they receive on account of those social ties (Granovetter, 1973; Burt, 1995; Rosenthal, 1996). The benefits that individuals or groups of individuals receive on account of their social ties is typically referred to by the term social capital. We were interested in testing this hypothesis, that social network structure relates to social capital, in the context of online social networks.

Our primary viewpoint towards the problem was through the lens of structural holes theory (Burt, 1995). A typical feature of social networks is that they consist of dense clusters linked by occasional bridge connections between the clusters. The "holes" in the network between these dense clusters of individuals who are not interacting are referred to as structural holes. Individuals within a cluster are likely to be of similar background due to homophily (Burt 2004). As a result, structural holes give us an indication of the diversity of individuals' contacts - those who act as bridges between clusters are likely to be in contact with diverse ties. Since this theory was our primary viewpoint in looking at the relationship between social network structure and social capital, we were specifically interested in testing whether, in an online social network (specifically, on Facebook), structural holes relates to social capital. Nevertheless, we decided to keep our investigation explorative, and therefore also examine the relationship between a few other social network metrics and social capital.

The second issue we investigate is whether individual differences affect the relationship between network structure and social capital, and how. Prior work has suggested that differences between individuals, such as in communication skill and self-esteem (Burke et al., 2011), influences how they obtain social capital from online social networks.

While numerous kinds of individual differences can affect the relationship between network structure and social capital, we are interested in observing whether we can detect more fundamental behavioural tendencies affecting this relationship. For this reason, we chose to examine the effect of personality traits, as characterized by the big five model of personality (Costa and McCrae, 1992).

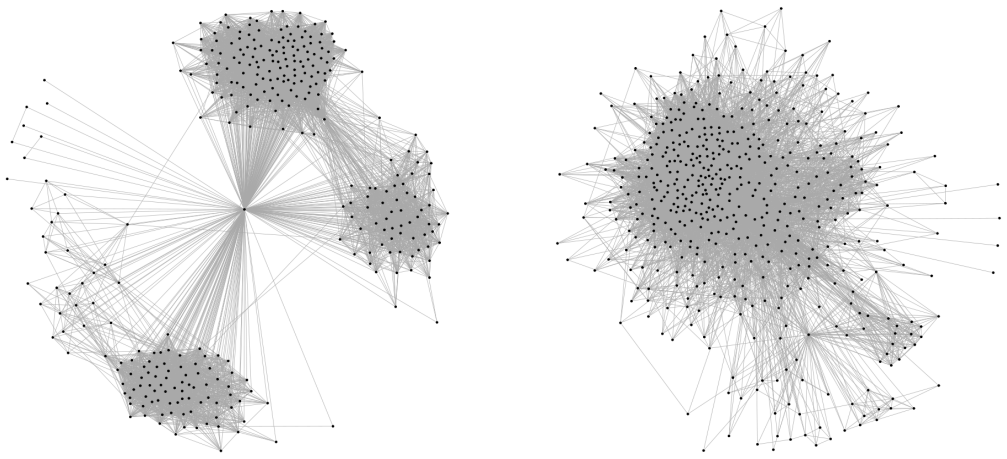

Figure 5.1. Examples of ego networks. In the center is the "ego" - a participant for whom we have collected data from their Facebook profile and through standard questionnaires. Around the ego are her "friends" in the Facebook social network. Also depicted are the friendship links between the individuals in the ego networks. The network on the right has lesser structural holes than the network on the left.

## 5.2 Main Article : Network, Personality and Social Capital

The paper is divided into five main sections. The first section introduces the problem we seek to examine. Section 2 provides a description of the data we collected for our study, including both network structure data and the social capital and personality scales. This section also describes of the network structure metrics we use in the analysis of the data provided in the subsequent chapter. Section 3 reports the results of our analysis of the metrics of social network structure, social capital, and personality traits. Section 4 discusses these results and Section 5 explains the limitations of the work.

Before we move on to the paper, it might be helpful for the reader to be introduced to the main network concepts and metrics we use in the paper. For a given participant, the ego-centric network around the participant, also referred to in the paper as an ego

network, contains the friends of the participant and the links between these friends (Figure 5.1).

As explained above in the introduction section of this chapter, a typical feature of social networks is that they consist clusters of dense connections linked by occasional bridge connections between the clusters (Burt, 1995). The "holes" in the network between these dense clusters or between individuals who are not interacting are referred to as structural holes (The contrast between the two networks in Figure 5.1 serves to illustrate this concept). The concept of structural holes is of interest in our current context since individuals who act as bridges between structural holes can benefit by having access to information and resources circulating in different clusters, and by acting as intermediaries between these clusters of people who are not directly interacting with each other.

Structural holes can also be thought of as equivalent to an opposite concept : Constraint (Burt, 2004). In a network with many structural holes the ego has greater control over the flow of information between his ties, since they form largely disconnected clusters. On the other hand in a network with few structural holes, the ego has limited control over the flow of information among all his ties. For example, if the ego in a network with fewer structural holes does something that is considered unacceptable in her community, this information can pass on to the rest of her ties and she has little control over the flow of the information. Hence individuals with networks having fewer structural holes are more "constrained".

The following are the metrics we use to quantify structural holes in our analysis of ego networks:

- *Effective Size:* The effective size is defined as

$$S = n - 2t/n$$

  where $n$ is the number of friends of the ego and $t$ is the total number of ties in the ego network not counting ties to the ego. Hence fewer the ties between the ego's friends, greater the effective size of the ego network.

- *Network Constraint:* An ego's brokerage opportunities are considered to be "constrained" if there exist alternative paths along which the information that she can broker between two individuals might travel, thus causing her to potentially lose those brokerage opportunities. Any contact *j* constrains the ego's brokerage opportunities to the extent that: (a) the ego has spent time and energy to form and maintain the tie with *j* (which she could have spent on other ties), and (b) *j* is a recipient of the time and energy spent by the other ties of the ego (Burt 1992). The constraint measure effectively captures the extent to which that is the case for each contact *j* of the ego *e* :

$$C = \sum_j [\ p_{ej} + \sum_q p_{eq} p_{qj}\ ]^2,\ j \text{ varies over friends of } e,\ q \text{ varies over friends of } j$$

Here $p_{ij}$ is the proportion of time that node *i* has spent on node *j* for any two nodes *i* and j in the ego network. When no specific metric of time spent is available or applicable, we can make the simplifying assumption that that an actor distributes his time equally over his contacts: if *i* is linked to *j*, then $p_{ij} = 1/d_i$, where $d_i$ is the number of ties of *i* within the ego network. If *i* and *j* are not linked, $p_{ij} = 0$. Constraint on a person is high if the person has few contacts (small network) and those contacts are strongly connected to one another, either directly (as in a dense network), or through a central, mutual contact (as in a hierarchical network).High constraint networks exhibit fewer structural holes, while low constraint networks exhibit more structural holes.

- *Betweenness Centrality:* Betweenness centrality captures the relative importance of an ego in the quick transmission of information within the ego network. Proposed by Freeman (1977), it measures the extent to which a person brokers indirect connections between all other people in the network. The betweenness centrality of an individual node "*v*" in a network is defined by the following formula:

$$\sum_{s,t} (\sigma_{st}(v)/\sigma_{st}),\ s \neq v \neq t$$

where $\sigma_{st}$ is the total number of shortest paths from node *s* to node *t* and $\sigma_{st}(v)$ is the number of those paths that pass through *v*. In addition to being a useful indicator of structural holes, the betweenness centrality of different ego networks can also serve as a reasonable approximation for the betweenness centrality of

these egos in the social network considered in its entirety (Everett & Borgatti, 2005).

In addition to measures of structural holes, we include the following network structure metrics in our analysis :

- *Degree Centrality:* We use the graph theoretic terminology of "degree" to refer to the number of connections an individual has. In our context, the degree of an individual is simply the number of Facebook friends the individual has. This number reflects the number of contacts the ego potentially has to access information and support.
- *Isolated Friends:* The number of isolated friends of an individual is the number of her friends with whom she has no other common contact. In the context, of Facebook, isolated friends are also likely to be online-only ties (people whom the ego has not met face-to-face). Hence such ties are likely to be weak ties, but by the same virtue likely to open access to new information to the ego.
- *Density:* The ratio of the number of links and the total number of possible links in an ego network. Density serves as a good indicator of how tightly knit the whole network of ties around the ego is.
- *Transitivity:* The probability that any two friends of an individual in the ego network are in turn friends. Transitivity serves as good indicator of the presence of tightly knit clusters or communities within the ego network.

*Venkatanathan, J., Karapanos, E., Kostakos, V., Goncalves, J. (2012). Network, Personality and Social Capital. In proceedings of ACM Web Science, Evanston, USA, p. 326-329. © 2012 ACM, Inc. Reprinted by permission. doi: 10.1145/2380718.2380760.*

# Network, Personality and Social Capital

Jayant Venkatanathan, Evangelos Karapanos
Madeira Interactive Technologies Institute
University of Madeira
Funchal, Portugal
{vjayant, e.karapanos}@m-iti.org

Vassilis Kostakos, Jorge Gonçalves
Department of Computer Science and Engineering
University of Oulu
Oulu, Finland
{vassilis, jgoncalv}@ee.oulu.fi

## ABSTRACT

We present a study on the relationship between social network structure on Facebook and social capital, and how this relationship is moderated by personality traits. The findings suggest that one's number of friends does not necessarily have an effect on the amount of bridging social capital. Conversely, the extent of structural holes and isolated friends in the network have an effect on bridging social capital. In addition, individuals low on agreeableness benefit more from isolated friends in terms of bridging social capital. In terms of bonding social capital, introverts benefit more from networks with higher transitivity. Women overall report higher bonding social capital, but there are no significant gender differences when it comes to leveraging one's network structure for bridging or bonding social capital.

## Categories and Subject Descriptors

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

## General Terms

Human Factors.

## Keywords

Social networks analysis, ego networks, social capital, personality traits.

## 1. INTRODUCTION

In our work we seek to understand the extent to which modern social networking systems can help individuals in their daily interactions through various means and support. In particular, we present a study that examines the relationship between individuals' social network structure on Facebook and social capital.

Previous studies [2, 10] found the structure of the networks around individuals predicted their success in an organization. This lead to the hypothesis that on account of their social ties and the structure of the network of these ties, certain individuals had access to more and a broader range of resources. In other words, these individuals had access to more and a broader range of social capital. The effect of the access to this social capital was therefore manifested in the overall outcome of higher success levels within the organization.

Social capital is the value of relationships between individuals and groups, and the resources and support that an individual has access to on account of his or her social ties. Social capital today is generally described using the constructs of *bridging* and *bonding* social capital [8]. Bridging social capital refers to the social capital created from bonds across individuals of different backgrounds. While these ties may lack in depth, they provide individuals with a broader horizon and open opportunities for new resources and information. Conversely, bonding social capital is created in bonds within individuals of a closed group such as family and close friends. These ties provide substantial and strong emotional support.

Previous work on network structure and social capital highlight two issues. First, it is not clear which kind of social capital, bridging or bonding, is associated with the network structure around individuals and their success in the organization. At the time, researchers drew from concepts such as the strength of weak ties [7], arguing that success was largely the result of bridging social capital, as we refer to the term today. While network structure might influence bridging social capital, one can expect that bonding support within these organizations might also have influenced the outcome of these individuals' success. Hence this raises the following question: *Can the structure of social ties around an individual independently help us predict the constructs of bridging social capital and bonding social capital?*

Second, individual differences can play a role in how positional advantages offered by network structure are leveraged. For example, certain individuals may have no inhibitions in approaching a distant tie for help in obtaining a job, while others might not be comfortable doing so. Thus, opportunities to leverage network ties and structure need not necessarily turn into social capital. Therefore, in addition to understanding how network structure influences social capital, it becomes important to understand: *how do individual differences in personality affect the leveraging of network structure for social capital*?

To answer these two questions we take advantage of the large-scale and granular availability of social network data on Facebook. Recent work has studied how social capital is leveraged on Facebook through the types of activities individuals engage in [3, 4, 12]. However, no work, to our knowledge, has examined how the network structure of social ties around individuals, as captured by online social networks, influence bridging and bonding social capital.


*WebSci 2012*, June 22–24, 2012, Evanston, IL, USA.

Past research examining the relationship between network structure and outcomes of social capital has typically made the implicit assumption that the direction of causality is from the former to the latter [Eg. 2, 6]. While these two variables are likely to influence each other to an extent, we assume that the dominant direction of causality is from network structure to social capital.

# 2. STUDY

Our study examines the effect of network structure in Facebook on social capital. We collected data from participants who gave us access to their list of friends on Facebook. From this data we were able to reconstruct their social network and calculate a number of metrics regarding their position in the network. Each participant also responded to standardized questionnaires of social capital [11] and the big five personality traits [9].

## 2.1 Participants

Participants were recruited through announcements and email lists in a university in Portugal and on social media targeting English speakers. Participants were also asked to rate their fluency in English. A total of 97 individuals (59 male) from 11 countries successfully completed the survey with an average age of 28 years old (sd=5.0). Participants had on average 303 friends (SD=178, max=875, min=9). Participants with less than 20 friends (N=2) were removed as they exhibited very little network structure and were likely to bias the results.

## 2.2 Measures

### *2.2.1 Network Analysis Metrics*

A typical feature of social networks is that they consist clusters of dense connections linked by occasional bridge connections between the clusters. The "holes" in the network between these dense clusters or between individuals who are not interacting are referred to as structural holes [1]. The concept of structural holes is of interest in our current context since individuals who act as bridges between structural holes can benefit by having access to information and resources circulating in different clusters, and by acting as intermediaries between these clusters of people who are not directly interacting with each other. Structural holes were quantified through the use of the following metrics:

- *Effective Size* captures the relationship between number of friends and number of ties between them in the ego network. The fewer the ties between the ego's friends, the greater the effective size of the ego network (for the exact definition the reader may refer to [2]).
- *Constraint* is high in a small network of contacts who are close to one another, or strongly tied to one central contact. High constraint networks exhibit fewer structural holes while low constraint networks exhibit more structural holes [2].
- *Betweenness centrality* captures the relative importance of an ego in the quick transmission of information within the ego network [5].

In addition we examine the following metrics in relation to social capital:

- *Degree centrality*: The number of friends in the ego network.
- *Isolated friends*: The number of friends in the ego network with no other common friend with the ego.
- *Transitivity* : The probability that any two friends of an individual in the ego network are in turn friends.
- *Density* : The ratio of the number of links and the total number of possible links in an ego network.

### *2.2.2 Social Capital*

Bridging and bonding social capital was measured with an adapted version of Williams' (2006) Internet Social Capital scales [11], consisting of six items for bridging social capital (Cronbach's alpha= 0.581, items 1, 2, 4, 7, 8 and 10 of the original scale; examples:"I am willing to spend time to support general community activities" and "Interacting with people reminds me that everyone in the world is connected") and five items for bonding social capital (alpha= 0.654, items 1, 2, 3, 8 and 10; examples: "There is someone I can turn to for advice about making very important decisions" and "There is no one that I feel comfortable talking to about intimate personal problems"(reversed)).

### *2.2.3 Personality Traits*

Personal traits were measured with the 10 item questionnaire of the big five inventory (10-BFI) [9]. It consists of two items for each of the five personality traits:

- *Extraversion* refers to the tendency for the individual to be outgoing and sociable (alpha= 0.519).
- *Neuroticism* refers to the tendency to experience anxiety and negative emotions (alpha= 0.649).
- *Conscientiousness* is the extent to which an individual is orderly, self-disciplined and strives for achievement (alpha= 0.57).
- Individuals high on *Agreeableness* are socially flexible, trusting and adjusting. (alpha= 0.045). The alpha value for this trait is unusably low. Further examination showed that participants uniformly rated themselves very high on one of the two items. Hence that item was dropped, leaving us with a single item for this trait. The item used in the analysis for the agreeableness trait is "I see myself as someone who tends to find fault with others" (reversed).
- *Openness to experience*, or simply *openness*, refers to overall curiosity, and artistic and scientific creativity (alpha=-0.086). The low alpha value for this scale similarly makes it unusable. As in the case of agreeableness, participants uniformly rated themselves high on one of the items, and this item was dropped. The single item used for the analysis of openness is "I see myself as someone who has few artistic interests" (reversed).

Most of the scales show alpha values only on the border of acceptability. Since we use single items to measure agreeableness and openness, we must interpret the results involving these traits with caution. Accounting for fluency in English and country had no effect on the reliability of any of the scales, hence these variables were subsequently discarded. Before proceeding further with analysis, all participants' scale ratings were converted to normalized z-scores. Degree, betweenness and constraint had heavy-tailed distributions and hence were converted to log-scale.

## 3. RESULTS

An independent samples t-test showed that females reported significantly higher bonding social capital than males($t(93) = -2.36$, $p<0.05$; Males: mean -0.21, sd 0.95; Females: 0.28, sd 1.02). There was no significant effect of gender on bridging social capital ($p>0.1$).

### 3.1 Network structure

Regression analysis showed that there was a significant effect of betweenness on bridging social capital ($t(93) = 2.0$, $b=1.17$, $p<0.05$, model adjusted r-sq = 0.03) and a marginally significant effect of constraint on bridging social capital ($t(93) = -1.88$, $b=-.34$, $p<0.1$, model adjusted r-sq=0.03). We also found a significant effect of the number of isolated friends on bridging social capital ($t(93)=2.81$, $b=0.31$, $p<0.001$, model adjusted r-sq=0.07). These are shown in Figure 1 and 2. We found no significant effect of degree on bridging social capital.

Additional examination with personality traits showed a marginally significant interaction of conscientiousness with constraint ($t(91)=-1.87$, $b=-.41$, $p<0.1$, model adjusted r-sq=0.04). In particular, the inverse relationship between constraint and bridging social capital was stronger among individuals with higher conscientiousness. There was also a significant interaction of agreeableness with the number of isolated friends in predicting social capital ( $t(91) = -1.99$, $b=-0.20$, $p<0.05$, model adjusted r-sq=0.09), and the positive relation between bridging social capital and the number of isolated friends was stronger among those with low agreeableness (Figure 3). There was no significant effect of density or transitivity and no significant interaction between any network metric and gender in predicting bridging social capital ($p>0.1$).

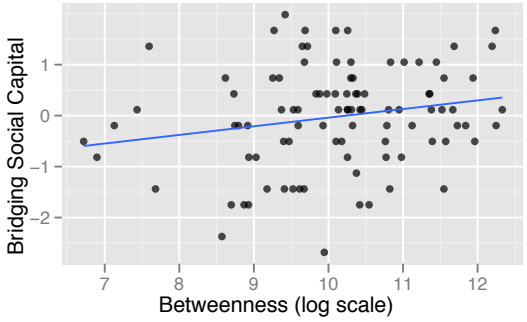


Figure 1. Betweenness vs Bridging social capital

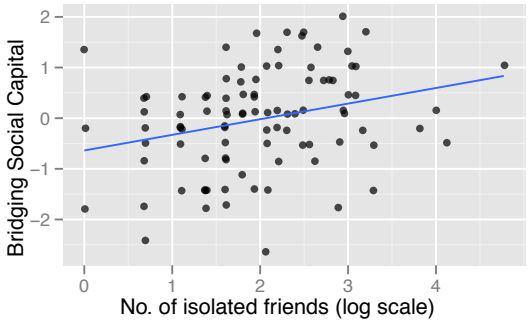


Figure 2. Isolated Friends vs Bridging social capital

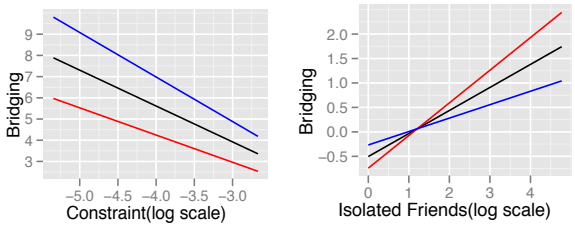


Figure 3: Left: Constraint Vs Bridging. Red - low conscientiousness (z-score= -1). Blue - high conscientiousness (z-score= +1).
Right: Isolated Friends Vs Bridging. Red - low agreeableness (z-score= -1). Blue - high agreeableness (z-score= +1)

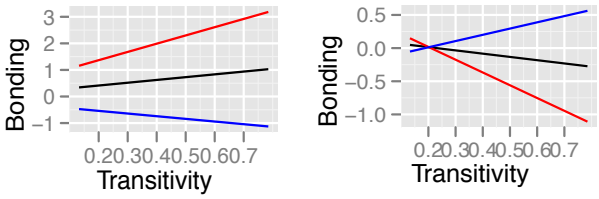


Figure 4: Transitivity vs Bonding.
Left: Red - low extraversion (z-score= -1). Blue - high extraversion (z-score= +1).
Right: Red - low openness (z-score= -1) . Blue - high openness (z-score= +1)

Further analysis showed that effective size had a marginally significant positive effect on bonding social capital ($t(93)=1.71$, $b=0.001$, $p<0.1$, model adjusted r-sq=0.02), and so did degree ($t(93) = 1.93$, $b=0.35$, $p<0.1$, model adjusted r-q = 0.03). The number of isolated friends had no significant effect on bonding social capital. Extraversion showed a significant interaction with transitivity in predicting bonding social capital ($t(91)= -2.92$, $b=-2.05$, $p<0.01$, model adjusted r-sq=0.08). In particular, introverts with higher network transitivity had higher levels of bonding social capital. In addition, openness showed a marginally significant interaction with clustering in predicting bonding social capital ($t(91)=1.90$, $b=1.43$, $p<0.1$, model adjusted r-sq=0.03). These interactions are shown in Figure 4. There was no significant effect of density and no significant interaction between any network metric and gender in predicting bonding social capital ($p>0.1$).

## 4. DISCUSSION

Overall, the study found network structure affecting both bridging and bonding social capital. While women reported higher bonding social capital, the findings show no evidence of gender differences when it comes to leveraging one's network structure for bridging or bonding social capital.

The findings revealed a positive effect of the extent of structural holes (measured by betweenness and constraint) in individuals' ego networks and bridging social capital. This is in agreement with the arguments put forward in prior literature to explain the effect of structural holes with success in organizational networks [2, 10].

Moreover, this positive effect of structural holes on bridging social capital was higher for conscientious individuals. An interpretation for this is that since conscientious individuals are self-disciplined and strive for achievement, they are better able to leverage the diversity of their network facilitated by higher structural holes, to obtain bridging social capital. The positive effect of structural holes on bonding social capital can be explained by the fact that individuals having networks with more structural holes are likely to have access to diverse ties for bonding needs who can provide different perspectives to a problem, or address distinct communication needs. However, these two results were only marginally significant, and hence should be treated with caution.

The positive effect of the number of isolated friends on bridging social capital confirms that such ties are likely to open up opportunities for new information and ideas. Interestingly, less agreeable individuals were likely to obtain higher bridging social capital from isolated friends. As expected, the number of isolated ties had no effect on bonding social capital, as these ties are likely to be weak ties and hence less likely to be a source of bonding support.

The number of friends of an individual had no effect on bridging social capital. This might suggest that merely increasing the number of friends does not lead to an increase in bridging social capital, unless that increase is through the addition of individuals from diverse backgrounds or communities (which is reflected in the measures of structural holes and isolated friends). There was a positive effect of the number of friends on bonding social capital, but this was only marginally significant.

Finally, introverts benefitted in terms of bonding social capital from higher transitivity. Since high transitivity networks consist of more closely knit clusters, this might suggest that introverts are better able to tap from closely knit networks for bonding needs.

## 5. CONCLUSION AND LIMITATIONS

This paper raised two questions: a) does network structure on Facebook predict social capital, and if so, b) is this relationship moderated by personality differences? Overall, the study suggests that the number of friends does not necessarily translate to bridging social capital, but the extent of structural holes and isolated friends in the network, along with personality, affect bridging social capital. In addition, introverts benefit in terms of bonding social capital from networks with higher transitivity.

One has to be cautious though in generalizing these results given the limits of our sample. First, due to the limited sample size, we were unable to inquire into higher order interaction effects, such as the extent to which individuals that are high on conscientiousness but low on extraversion are able to leverage their network structure for social capital. Secondly, a methodological drawback of the study is that participants were self-selected as they responded to an online survey call. This might have lead to a possible non-response bias in our sample, whereby the sample of Facebook users who chose not to respond to our announcements for the study might have shown an overall difference from our participants in terms of network structure, personality or social capital. Last, some of the scales showed very low reliability scores. While small item scales have the benefit of lower participant fatigue, low reliability can be expected for questionnaire scales containing only few items per construct, such as the personality traits questionnaire used in this study. Hence, these results need to be interpreted with caution.

## 6. ACKNOWLEDGEMENTS

This work was funded by the Portuguese Foundation for Science and Technology (FCT) grant CMU-PT/SE/ 0028/2008 (Web Security and Privacy). Additional support was provided by the Academy of Finland and TEKES.

## 7. REFERENCES

1. Burt, R.S. (2004). Structural holes and good ideas. American Journal of Sociology 110, 349-399.
2. Burt, R. S. (1995). Structural Holes: The Social Structure of Competition. Harvard University Press.
3. Burke, M., Kraut, R. & Marlow, C. (2011). Social capital on Facebook: Differentiating uses and users. Conference on Human Factors in Computing Systems CHI 2011, 571-580, ACM.
4. Ellison, N., Steinfield, C., & Lampe, C. (2007). The Benefits of Facebook "Friends:" Social Capital and College Students' Use of Online Social Network Sites. Journal of Computer-Mediated Communication, 12(4), 143-1168.
5. Everett, M. and Borgatti, S. P. (2005). Ego network betweenness. Social networks (Soc. networks) 27, 1 (2005), 31–38.
6. Granovetter, M. (2005). The impact of social structure on economic outcomes. Journal of Economic Perspectives 19, 1, 33–50.
7. Granovetter, M. (1973). The strength of weak ties. American Journal of Sociology 78: 1360–1380.
8. Putnam, R. (2000). Bowling alone: the collapse and revival of American community. New York: Simon and Schuster.
9. Rammstedt, B., and John, O.P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. Journal of Research in Personality 41, 203–212.
10. Rosenthal, E. A. (1996). Social Networks and Team Performance. Ph.D. Dissertation, Graduate School of Business, University of Chicago.
11. Williams, D. (2006). On and Off the 'Net: Scales for Social Capital in an Online Era. Journal of Computer- Mediated Communication, 11(2), article 11.
12. Yoder, C. and Stutzman, F. (2011). Identifying Social Capital in the Facebook Interface. Conference on Human Factors in Computing Systems CHI 2011, ACM.229.

## 5.3 Critical review of the work and conclusion

***A note on the results and the context of the direction of work***

While we did touch upon the limitations of this work in section 5 of the paper, here we would like to dwell a little more on this issue. This work explores the use of ego-centric networks as a source of data to help us understand social capital in the context of online social networks. Our goal here was to look into these phenomena in an explorative manner, so as to get a feel both for how these phenomena might be related, and for the potential of the methodology in helping us understand these phenomena. While our examination of the relationship between social capital and network structure (particularly structural holes) was motivated from theory in the social sciences (eg. Burt, 1995), and our results in the context of Facebook were in agreement with the prediction of this theory, our examination of the influence of personality on this relationship was explorative. In other words, we had no clear a priori hypotheses on how personality traits might affect the relationship between network structure and social capital, and instead we looked to our collected data to get a sense of this. Based on our data, we discussed the possible interpretations for how personality might be influencing the way in which individuals translate network structure into social capital (Section 5 of the paper), and also presented some graphs that illustrate the effect of personality in our data (Figure 4 of the paper). We consider these results, particularly on the effects of personality, not as concrete facts but rather as provisional insights that one can use to direct subsequent explorations and hypotheses.

Having elucidated the limitations of this stage of the work, we would like to explain the context of the current activity within the research community in which we see the value of this direction of work. The question of how the internet and online social networks affect our well being and social capital is a complex issue. Some early studies suggested that internet use is detrimental to our well being and led to declines in communication with family members in the household, the size of their social circle, and increases in depression and loneliness (Kraut et al., 1998). Yet other studies suggested that internet use had a positive effect on social capital by extending existing levels of face-to-face contact (Wellman et al., 2001).

Over the years, one of the primary ways in which researchers have sought to overcome the limitations of attempts to draw sweeping conclusions on the relationship between the use of online social networks and social capital is by carrying out increasingly finer grained analyses, such as by separately considering the different uses of online social networks (Burke et al., 2010), the different user interface elements in online social networks (Yoder and Stutzman, 2011) and even differences between individuals in terms their personality or skills (eg. Burke et al., 2011). This is akin to the strategy of "unbundling" of the various features and the various uses of social networks, a term used by Smock et al. (2011) in their study of the uses and gratifications of Facebook. The strategy of "unbundling" entails increasingly granular analyses that differentiate to the extent possible between contexts, users, uses, and the like to probe deeper into phenomena and resolve inconsistencies in findings. In our view, this is a natural and worthwhile direction of analyses that leads to a more nuanced understanding of the complex relationship between online social network use and social capital.

Our current approach to understanding the impact of SNS on social capital is by considering social network structure, an aspect of SNS that Human-Computer Interaction (HCI) research on social systems has not fully adopted yet. Social Networks Analysis (SNA) has a long history in the social sciences, and there is a body of work on network structure grounded in theories of how individuals and groups interact and affect each other (e.g., Berkowitz, 1982; Granovetter, 1983). As a result, SNA lends itself naturally to the validation of these theories of human interaction, which makes it valuable to help us make sense of the underlying mechanisms that govern these interactions. Therefore, this direction of investigation can ultimately contribute towards an integrated understanding of these phenomena based on theoretical foundations, and is complementary to the strategy of unbundling to further our understanding of social capital and SNS.

#### *Conclusion*

In this work on the relationship between online social network structure and social capital we sought to test some basic ideas on the role of network structure based on traditional social science in the context of online social networks, and explore the role of personality traits, based on the big 5 model, on this relationship. Our results on the relationship between network structure and social capital, based on network structure

data of a group of participants from the Facebook social network, are in alignment with the predictions of prior theory in social science, in particular the theory of structural holes. Further analyses on the effect of personality traits provide a number of provisional insights on how different individuals tap into their network for social capital. Following this, we explain why our work is a step in a new direction of investigation within HCI research on the social capital outcomes resulting from SNS use.

**Chapter References**

The following are the references cited in this chapter outside of the embedded paper.

Berkowitz, S. D. (1982). An introduction to structural analysis: The network approach to social research. Toronto: Butterworth.

Burke, M., Marlow, C. and Lento, T. (2010). Social network activity and social well-being. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI). ACM,1909–1912.

Burke, M., Kraut, R. and Marlow, C. (2011). Social capital on Facebook: Differentiating uses and users. In Proceedings of the Conference on Human Factors in Computing Systems (CHI). ACM, 571–580.

Burt, R. S. (1995). Structural Holes: The Social Structure of Competition. Harvard University Press.

Burt, R. S. (2004). Structural holes and good ideas. American Journal of Sociology 110, 349–399.

Costa, P. T. ‚Jr. & McCrae, R. R. (1992). Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) manual. Odessa, FL: Psychological Assessment Resources.

Everett, M. and Borgatti, S. P. (2005). Ego network betweenness. Social networks (Soc. networks) 27, 1 (2005), 31–38.

Freeman, L. C. (1979). Centrality in social networks: Conceptual clarification. Social Networks, 1(3), 215–239.

Granovetter, M. (1973). The strength of weak ties. American Journal of Sociology 78, 6, 1360–1380.

Heider, F. (1958). The Psychology of Interpersonal Relations. John Wiley and Sons.

Kraut, R., Lundmark, V., Patterson, M., Kiesler, S., Mukopadhyay, T., & Scherlis, W. (1998). Internet paradox: A social technology that reduces social involvement and psychological well-being? American Psychologist, 53(9), 1017-1031.

Rosenthal, E. A. (1996). Social Networks and Team Performance. Ph.D. Dissertation, Graduate School of Business, University of Chicago.

Smock, A. D., Ellison, N. B., Lampe, C. and Wohn, D. Y. (2011). Facebook as a toolkit: A uses and gratification approach to unbundling feature use. Computers in Human Behavior 27, 6, 2322–2329.

Wellman, B., Haase, A. Q., Witte, J., & Hampton, K. (2001). Does the Internet increase, decrease, or supplement social capital? Social networks, participation, and community commitment. American behavioral scientist, 45(3), 436-455

Yoder, C., & Stutzman, F. (2011). Identifying social capital in the Facebook interface. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 585-588). ACM.

# 6. Understanding Empathy Through Network Structure

## 6.1 Introduction

This chapter presents paper (d) "*A Network Science Approach to Modelling and Predicting Empathy*", presented at the International Workshop on Web Behavior Analytics (2013) and published in the adjunct proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. Here we take the work from the previous chapter one step further by considering a human ability that is fundamental to successful relationships, empathy, and the relation of this ability to network structure and social capital. Empathy is the ability to identify with and understand another's situation, feelings and motives (Preece, 1999). Unlike other universal behavioural tendencies considered in this thesis such as reciprocity, homophily and personality traits, empathy does not represent a fixed behaviour pattern. Nevertheless, like these other universal behavioural tendencies, empathy is a psychological predeterminer of behaviour, and has a significant influence on our interaction behaviour. It is for this reason that empathy is important in the context of this thesis.

In this chapter we adopt a networks analysis perspective to study empathy. Like in the study presented in the previous chapter, we use data collected from participants of the Facebook social network for our analyses. However, unlike our treatment of personality traits in the previous chapter, we have clear hypotheses on how we might expect empathy to relate with network structure and social capital, grounded on prior literature (Eg. Boisjoly et al., 2006; Galinsky, 2002; Wölfer et al., 2012), and therefore this work is a more focussed examination of these hypotheses based on our collected data. Overall, the results from our data are in alignment with the predictions we draw from prior literature, suggesting that empathy is indeed related to online network structure and therefore we might be able to use network structure as a lens through which we can study aspects of this fundamental human ability in the context of online social networks.

## 6.2 Main Article : A Network Science Approach to Modelling and Predicting Empathy

The paper is divided into six main sections. The first two sections provide the motivation and grounding of the research question, leading into the specific hypotheses we set out to examine in the study. The third section describes the study and the key measures of empathy, social capital and network structure that were measured. The fourth section reports the analyses we performed to test our hypotheses. The fifth section discusses the results of the analyses, in terms of understanding empathy from network structure (section V. *A*) and in terms of predicting empathy from network structure (section V. *B*). We would like to downplay this second discussion of predicting empathy from network structure, as we have only performed correlational analyses that indicate a relationship between empathy and network structure at the population level, but do not prove that we can predict empathy at the level of individuals. In the final section we discuss the limitations of the study.

*Venkatanathan, J., Karapanos, E., Kostakos, V., Goncalves, J. (2013). A Network Science Approach to Modelling and Predicting Empathy. In adjunct proceedings ASONAM 2013, International Workshop on Web Behavior Analytics. Niagara Falls, Canada.*
 *doi: 10.1145/2492517.2500295*

# A Network Science Approach to Modelling and Predicting Empathy

Jayant Venkatanathan, Evangelos Karapanos
Madeira Interactive Technologies Institute
University of Madeira
Campus da Penteada, Funchal, Portugal.
{vjayant, e.karapanos}@m-iti.org

Vassilis Kostakos, Jorge Gonçalves
Department of Computer Science and Engineering
University of Oulu
Oulu, Finland
{vassilis, jgoncalv}@ee.oulu.fi

***Abstract* — In this paper we adopt a network science approach to investigate empathy and its implications for online social networks. We demonstrate that empathy is closely linked to social capital - the findings suggest that individuals higher on cognitive empathic skill are overall likely to report both higher bridging and higher bonding social capital. On the other hand, attributes of network structure around the individual, quantified through networks analysis metrics, were related to cognitive empathy. Further, an examination of the interplay between network structure, social capital and empathy suggests that empathy facilitates the relation between network structure and social capital previously reported in literature. We discuss the implications of our findings for the understanding of empathy in the context of online social networks and for the design of these systems.**

***Keywords* — *Empathy; social capital; ego networks; online communities***

## I. Introduction

Individuals are increasingly establishing social capital by turning to online social networks for support - social, emotional, psychological (Eg [16]). An important research challenge in this context is to develop an operational understanding of how social capital can help foster online communities. In this paper we focus on one aspect of social capital, empathy, and its relationship with social capital and social network structure. This understanding can ultimately used to enhance and foster online communities.

Empathy is an important trait that enables us to "tune in" to others' feelings and thoughts. It can be described as the ability to feel or imagine another person's emotional experience [13]. Empathy allows us to understand the intentions of others, predict their behaviour, and experience an emotion triggered by their emotion [1]. Thus, the ability to empathise enables us to interact effectively with others, both face to face and online, and is fundamental to successful human relationships.

One useful approach to understanding and drawing insights into social interactions is social networks analysis. While on the one hand the representation of relationship ties as links in a network is a simplification, it is this very simplification that makes it valuable for the population level analysis of personality traits. Given that empathy and social interactions are closely tied, can the "fingerprints" of empathy then be found in social network structure? If so, then networks analysis can be used as a lens with which to study empathy, to the extent to which empathic skill is tied to social interactions. Thus the answer to this question can have implications for understanding empathy and also for the design of systems that foster empathic relating between users.

Motivated by this basic question, in this work we adopt a network science perspective to investigate how online social network structure can help us understand and predict empathy. To achieve this we take advantage of the large-scale and granular availability of social network data on Facebook.

The following are the contributions of the paper : (1) We show that empathic ability and social capital are closely related. (2) Through sociometric analysis we find a link between an individual's social network structure and empathic ability. (3) We demonstrate through mediation analysis that this link facilitates the previously reported [19] link between network structure and social capital. (4) Finally, we draw insights from our findings on how design can foster communities, how to target advertising, and how to enhance affective computing applications.

## II. Background

Despite the strong link between social interaction and empathy shown in literature, just one prior study has considered social network analysis as a proxy for studying empathy. Specifically, Wölfer et al. [21] recently showed that empathy is mirrored in the structure of social ties

This work was funded by the Portuguese Foundation for Science and Technology (FCT) grant CMU-PT/SE/ 0028/2008 (Web Security and Privacy). Additional support was provided by the Academy of Finland and TEKES.

among adolescents in German schools, as recorded through face-to-face interactions.

Here we extend this prior work to examine how empathic ability is reflected in the social network structure around individuals. To do this, we take advantage of the availability of social network data from a cohort of participants on the Facebook social network site. Facebook is typically used as a means of building and maintaining relationships involving those with whom users share "some common offline element" [3]. However, Facebook enables users to "convert latent and weak ties" [7] and is therefore particularly useful for developing bridging social capital [18]. Thus network structure on Facebook, while closely related to and impacted by the offline network, has it's own role and impact on individuals.

While little previous work has examined empathy directly with online social network structure, a number of previous studies suggest that empathy may affect network structure. Interaction between individuals of diverse backgrounds, such as diverse ethnicity [2] can lead to increased empathy. Conversely, those with increased empathy are more comfortable with individuals of diverse background, and for example, reduce out-group stereotyping [9]. While prior work has not directly linked empathy to network structure, it does suggest that structural holes [5] indicate diversity. Therefore, one way to address our research question is to investigate whether individuals whose networks contain relatively more structural holes are more empathic.

Empathic individuals are, by definition, likely to better understand others' needs and distress, and are thus likely to provide social support [21]. As a consequence of social reciprocity [10], empathic individuals are thus more likely to receive help from others. Another way to address our research question, therefore, is to establish whether individuals with higher empathy have increased social capital.

Prior work has also directly linked social network structure to social capital [19]. Given this finding, if empathy is indeed reflected in social network structure, then one may expect that empathy helps individuals exploit their network structure for social capital. For example, more empathic people may better translate potential resources in their network structure into social capital. In other words, we can seek to address our research question by investigating whether empathy moderates the relationship between network structure and social capital.

Finally, should we find evidence to support our assertions (i.e. links between empathy & structural holes, and empathy & social capital), then one might expect that the influence of network structure on social capital happens partly through empathy: network structure affects empathic skills, which in turn lead to higher social capital. Hence, we can further address our research question by attempting to establish whether empathy mediates the relationship between network structure and social capital.

## III. Study

93 participants (57 male; average age 28.2, sd 5.1) were recruited through online announcements and emails. Each participant gave us access to their list of friends, and the friendships between these friends, on Facebook, from which we were able to construct their social network and calculate a number of structural metrics regarding their position in the network. Participants had on average 315 friends (SD=172, max=875, min=50). In addition to providing us access to their Facebook social graph, each participant responded to standard questionnaires of empathy [15] and social capital [20].

### *A. Network Structure*

We use measures of structural holes to capture the diversity each participant's social network. A typical feature of social networks is that they consist of dense clusters linked by occasional bridge connections between the clusters. The "holes" in the network between these dense clusters of individuals who are not interacting are referred to as structural holes [5]. Individuals within a cluster are likely to be of similar background due to homophily [5]. Therefore, structural holes is of interest to us as those who act as bridges between clusters are exposed to diverse ties. Structural holes are quantified through the network constraint and betweenness centrality metrics :

- Constraint is high in a small network of contacts who are close to one another, or strongly tied to one central contact. High constraint networks exhibit fewer structural holes while low constraint networks exhibit more structural holes [4].
- Betweenness centrality captures the relative importance of an ego in the quick transmission of information within the ego network [8].

We also recorded the number of friends and the number of isolated friends (friends with whom the individual has no common friends) since these are known to be related with social capital [19].

### *B. Empathy*

Empathy was measured with the 8-item version [15] of the empathy quotient (EQ) scale [1]. A principal components analysis (varimax rotation, eigenvalue>1, loadings>0.6) on the items of the empathy scale revealed a three-factor structure, in agreement with Lawrence et al.'s [13]

validation of the original EQ scale. Following their labelling, these factors are: (1) Cognitive Empathy - the capacity to comprehend the emotions of others (items 2, 3 & 4 of the questionnaire from [15], example: "I am quick to spot when someone in a group is feeling awkward or uncomfortable"; eigenvalue=2.29, 28.65% explained variance, Cronbach $\alpha$=0.679), (2) Social Skills - knowing of how to behave in different social situations and the understanding of social norms (items 5 & 6, example: "I find it hard to know what to do in a social situation"; eigenvalue=1.47, 18.39% explained variance, Cronbach $\alpha$=0.52), and (3) Emotional reactivity - the extent to which individuals own emotional state is affected by other people's emotions (item 8 - "Other people often say that I am insensitive, though I don't always see why"; eigenvalue=1.07, explained variance 13.34%).

While it is not clear whether emotional reactivity is by itself a component of empathy, it is likely to tap into affective empathy (the capacity to experience others' emotions) [13]. The social skills factor shows a low reliability score, and emotional reactivity is measured by a single item. Therefore results following from these factors must be interpreted with care.

### C. *Social Capital*

Social capital is generally described using the constructs of bridging and bonding social capital [18]. Bridging social capital refers to the social capital created from bonds across individuals of different backgrounds. While these ties may lack in depth, they provide individuals with a broader horizon and open opportunities for new resources and information. Conversely, bonding social capital is created in bonds within individuals of a closed group such as family and close friends. These ties provide substantial and strong emotional support.

Bridging and bonding social capital were measured with an adapted version of Williams' [20] Internet Social Capital scales, consisting of six items for bridging social capital (Cronbach $\alpha$=0.581, example: "Interacting with people reminds me that everyone in the world is connected") and five items for bonding social capital (Cronbach $\alpha$=0.654, example: "There are several people I trust to help solve my problems"). Detailed information about the items used for bridging and bonding can be found in [19].

Before proceeding further with analysis, all participants' scale ratings and network metrics were converted to normalised z-scores. Degree, betweenness and constraint had heavy-tailed distributions and hence were converted to logarithmic scale.

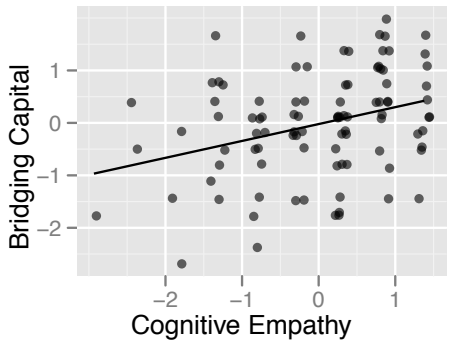


Fig 1. Cognitive empathy vs Bridging Social Capital (darker dots indicate overlapping points).

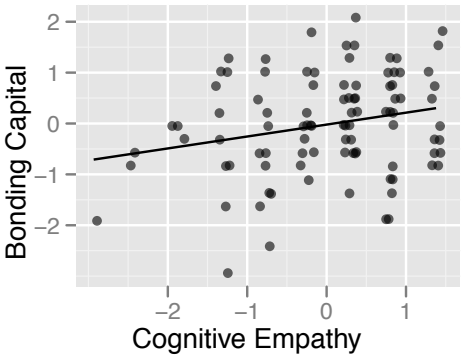


Fig 2. Cognitive empathy vs Bonding Social Capital

## IV. Results

Independent samples t-tests showed no significant gender differences in any of the empathy subscales ($p>0.05$). However, overall females reported significantly higher bonding social capital than males ($t(91)=-2.21$, $p<0.05$; Males: mean -0.18, sd=0.95; Females: 0.28, sd=1.02). There was no significant effect of gender on bridging social capital. While age was positively related to social skills ($b=0.047$, $t(91)=2.37$, $p<0.05$, r-sq=0.048), it was not significantly related with cognitive empathy or emotional reactivity, nor bridging or bonding social capital ($p>0.05$). However, since our participants largely comprised of young individuals, our results might not capture the true effect of age on these variables.

### A. *Social Capital and Empathy*

Next, we examined the relationship between the 3 factors of the empathy scale and social capital. Regression analysis showed that cognitive empathy had a significant positive relationship with both bridging social capital ($b=0.325$, $t(91)=3.275$, $p<0.01$, r-sq=0.106)) and bonding social

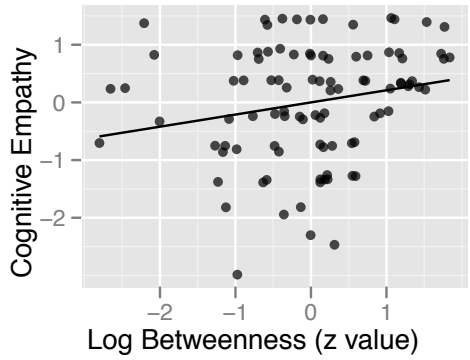


Fig 3. Betweenness (log scale) vs Cognitive Empathy

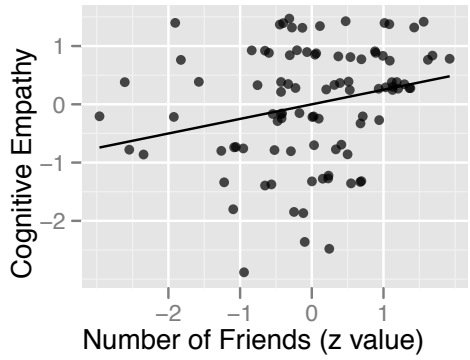


Fig 4. Number of Friends vs Cognitive Empathy

capital (b=0.223, t(91)=2.212, p<0.05, r-sq=0.051). These are shown in Figures 1 and 2. Social skills showed no significant relationship with either component of social capital.

Emotional reactivity showed a significant relationship with bonding social capital (b=0.377, t(91)=3.93, p<0.01, r-sq=0.145), but not with bridging social capital (p>0.05). Overall the results show that empathy is positively associated with both bridging social capital (cognitive empathy) and bonding social capital (cognitive empathy and emotional reactivity).

### *B. Network Structure and Empathy*

Measures of structural holes (betweenness and constraint) showed a significant relationship with cognitive empathy (Figure 3). For betweenness (log scale z scores) : b=0.214, t (91)=2.092, p<0.05, r-sq=0.046. This confirms that individuals with networks containing larger structural holes are likely to have higher cognitive empathy. However, betweenness (and other measures of structural holes) showed no significant relationship with either social skills or emotional reactivity (p>0.05).

The number of friends also showed a significant relationship with cognitive empathy (b=0.264, t(91)=2.393, p<0.05, r-sq=0.059) (Figure 4), but not with social skills or emotional reactivity (p>0.05). The number of isolated friends showed no significant relationship with either of the 3 factors (p>0.05).

Structural holes (measured by betweenness) also showed a significant relationship with bridging social capital (b=0.205, t(91) = 2.0, p<0.05, r-sq=0.042), and so did the number of isolated friends (b=0.283, t(91)=2.80, p<0.05, r-sq=0.08). We refer the reader to [19] for a complete discussion on the relationship between network structure and social capital. In this paper we focus on empathy and its relationship to these variables.

### *C. Role of Empathy in Social Capital – Network Structure Relationship*

Multiple regression analyses showed no significant interaction between any empathy factor and any of the network metrics in predicting bridging or bonding social capital (p>0.05). Thus our results suggest that that empathy does not play a moderating role in the relationship between network structure and social capital.

Finally, we examined whether empathy played a mediating role in the relationship between structural holes and bridging social capital. As suggested by Preacher and Hayes [17], especially for small samples, we conducted a bootstrap analysis for indirect effects to test for mediation. Based on 5000 bootstrap samples, the analysis found empathy to be a significant mediator in the relationship between betweenness and bridging social capital (confidence interval [0.0108, 0.1566], p<0.05, Data=0.0614, boot=0.0613, bias=-0.0002, SE=0.0353). This suggests that empathy is, in part, an intermediate variable in the translation of network structure to bridging social capital.

## V. Discussion

Our study set out to understand the relationship of empathy to social capital and social network structure. The ultimate goal of such work is towards drawing insights on fostering online communities. We find that empathy is related to social capital, and show how social network structure can be used to understand empathy. Below we discuss how these findings enhance our understanding of empathy, and how these insights can be ultimately lead to design for fostering online communities.

### A. Understanding Empathy

Our work shows meaningful trends in the relationship between online network structure and empathy. Thus a network science approach can provide a novel lens with which to study empathy. In particular, we find a consistent relationship between structural holes and empathy. While Wölfer et al. [21] report that in face to face networks between adolescents in classrooms those who are more central show higher empathy, we find similarly that the extent of structural holes is significantly related to empathy in more general online social networks of adults. The large-scale automated analysis available with online social networks make this an important opportunity for population-level analysis of this trait.

In addition, the relationship between empathy and bridging and bonding social capital suggest that it is an important ability that facilitates individuals to support and draw support from each other. There are two possible ways this can happen. First empathic individuals, by nature of inclination towards prosocial activities [21], increase the overall social capital of the community of ties around them, thus indirectly affecting their own social capital. Second, such individuals are likely to receive direct reciprocity [10] for their support, and thus experience higher social capital. By either mechanism, empathic individuals increase the social capital of the community, due to which this skill can facilitate community fostering.

The evidence for a mediation role of empathy in the relationship between network structure and social capital further reiterate the importance of this skill in communities. This result suggests that part of the translation of network structure to social capital is able to take place due to the effect of structural holes on empathic skill. While careful confirmation of the exact direction of causality will require longitudinal assessment, our findings suggest that these factors, to an extent, vary together.

### B. Informing Design

Empathy is known to be consistently related with prosocial activities [21]. Thus individuals with higher empathy are more disposed to help others in the network, and participate in overall community building. Our findings show that it is feasible to predict empathic ability through automated analysis of social network structure. If we can predict empathy using sociometric analysis, then we can identify those particular manifestations of human behaviour in a large network. Facebook is uniquely able to see a "macro" view of empathy across the network, and therefore can propose "interventions" that will influence social networks and communities in a number of ways.

One way in which the ability to predict empathy can be exploited is in the fostering of online communities. Individuals who are likely to possess aspects of empathic skill such as cognitive empathy, which are traceable in network structure, can be identified, and the support of these individuals can be leveraged. For example, in online support communities such as those for quitting smoking or for coping with depression, it can often happen that certain individuals are unable to get draw the support they require from the community, be it due to a difficulty to communicate on their part or a difficulty on the part of members in the community to understand their support needs. Such users might have better chances of response from members who are likely to have high cognitive empathic skills, and thus such members can be highlighted for these users to draw support from.

Predicting empathy can also be used to improve audience targeting for organisations such as those working on social causes. Particularly, being able to identity the different kinds of empathy in an individual can inform the design of targeted calls for support. For example, individuals with higher affective empathy might be likely to better identify with videos showing the people in distress, while for those higher on cognitive empathy it might be more important to highlight the background situation that is causing the distress for which the cause attempts to help.

Finally, there is a body of work on affective computing which attempts to identify the emotional states of the user, such as frustration, and thereby provide appropriate responses to reduce this frustration [12]. However, there has been skepticism about the "canned" response of affective computers [13]. One way to overcome the drawback of "canned" interventions in a social networking setting is to highlight the presence of empathic members who are experts in the network. Those with high cognitive empathy are likely to be more effective at communicating with other users, and therefore such a member is likely to better assist an individual who is identified to be stressed by the use of the application.

While the ideas presented above are of a speculative nature and concrete design requirements will require further maturation of this area of work, we have attempted to provide a glimpse into possibilities that can result from understanding empathy in a social network setting.

## VI. Limitations and Conclusion

This paper set out to study the relationship between empathy and network structure and social capital. We find that empathy is related to both the structure of individuals' networks and the social capital they report. A limitation of the study is the modest sample size drawn largely from a young population. In addition, it is important to recognise that techniques to gather data on individuals' empathy have

inherent limitations. Particularly, there are limitations in self report as it involves subjective assessment. We could alternatively use more objective tests such as the ability to recognise facial expressions [11] or even brain activity [6]. However, such tests do not directly measure empathy itself, but rather certain underlying mechanisms related to empathy. While the approach we adopted was most appropriate and feasible for this work, future work can consider different or multiple approaches to measuring empathy.

It is important to stress that empathy is not purely determined by network structure, but rather that the way the network structure evolves reflects and affects certain aspects of empathy. These traces of empathy can be detected in network structure over a macro view of the population. Clearly, other factors such as the nature of close and intimate relationships also affect empathy, which network structure might not capture. What our work shows is that network structure sufficiently reflects empathy to be detectable, and to that extent can be understood through social networks analysis. While we drew a number of findings from the current analysis, these have implications both for the understanding of this fundamental human trait and for the fostering of communities in online social networking for future work to explore.

## References


1. Baron-Cohen, S. & Wheelwright, S. (2004). The Empathy Quotient: an investigation of adults with Asperger's syndrome or high functioning autism, and normal sex differences. Journal of Autism & Developmental Disorders.
2. Boisjoly, J., Duncan, G.J., Kremer, M., Levy, D.M. and Eccles, J. (2006). "Empathy or Antipathy? The Impact of Diversity." American Economic Review, 96(5).
3. Boyd, D. M., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. Journal of Computer-Mediated Communication, 13(1).
4. Burt, R. S.(1995). Structural Holes: The Social Structure of Competition. Harvard University Press.
5. Burt, R.S. (2004). Structural holes and good ideas. American Journal of Sociology 110, 349-399.
6. Carr, L., Iacoboni, M., Dubeau, M. C., Mazziotta, J. C., & Lenzi, G. L. Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. Proceedings of the National Academy of Sciences of the United States of America PNAS (2003), vol. 100 no. 9: 5497–5502.
7. Ellison, N., Steinfield, C. & Lampe, C. (2011). Connection Strategies: Social capital implications of Facebook-enabled communication practices. New Media & Society.
8. Everett, M. and Borgatti, S. P. (2005). Ego network betweenness. Social networks (Soc. networks) 27, 1 , 31–38.
9. Galinsky, A. D. (2002). Creating and reducing intergroup conflict: The role of perspective-taking in affecting out-group evaluations. In M. A. Neale, E. A. Mannix, & H. Sondak (Eds.), Toward a phenomenology of groups and group membership (Vol. 4, pp. 85–113). Greenwich, CT: JAI.
10. Gouldner, A W. (1960): The norm of reciprocity: a preliminary statement. American Sociology Review. 25:161--78
11. Keltner, D., Ekman, P., Gonzaga, G.C., Beer, J. D., Richard J. (Ed), Scherer, K.R. (Ed), Goldsmith, H.H. (Ed). (2003). Handbook of affective sciences. Series in affective science., (pp. 415-432). New York, NY, US: Oxford University Press, xvii, 1199 pp.
12. Klein, J., Moon, Y. & Picard, R. (2002). This Computer Responds to User Frustration: Theory, Design & Results. Interacting with Computers 14: 119-140.
13. Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., David, A. S. (2004). Measuring empathy: reliability and validity of the Empathy Quotient. Psychological Medicine, 34, 911–924.
14. Lindgaard, G. (2004). Adventurers versus nit-pickers on affective computing. Interacting with Computers, 16, 723–728.
15. Loewen, P.J., Lyle, G. & Nachshen, J.S. (2009). An eight-item form of the Empathy Quotient (EQ) and an application to charitable giving. Retrieved from crcee.umontreal.ca/pdf/Eight%20Question%20ES_final.pdf
16. Pfeil, U., and Zaphiris, P. (2007). Patterns of empathy in online communication. Proc. CHI 2007, 919-928. ACM.
17. Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. Behavior Research Methods, Instruments, & Computers, 36, 717–731.
18. Putnam, R. (2000). Bowling alone: The collapse and revival of American community. New York: Simon and Schuster.
19. Venkatanathan, J., Karapanos, E., Kostakos, V., Goncalves, J. (2012). Network, Personality and Social Capital. In proceedings of ACM Web Science.
20. Williams, D. (2006). On and Off the 'Net: Scales for Social Capital in an Online Era. Journal of Computer-Mediated Communication, 11(2), article 11.
21. Wölfer, R., Cortina, K.S. & Baumert, J. (2012). Embeddedness and empathy: How the social network shapes adolescents' social understanding. Journal of Adolescence 35, 1295–1305.

## 6.3 Critical Review of the work and conclusion

### *A further note on limitations*

We would like to touch upon the issue of the effect size of our results in this work. Although the correlations between empathy and the network structure variables were statistically significant, the correlations were only modest. For example, the correlation between cognitive empathy and betweenness was 0.214 and the correlation between cognitive empathy and the number of friends was 0.264. This is not surprising as there are likely to be numerous aspects of empathy, such as the nature of family relationships, that network structure cannot capture. Similarly, there are likely to be numerous factors other than empathy that affect network structure (and social capital). Therefore it is expected that the magnitude of correlation between these variables is low even if these correlations represent true effects between the variables.

Further, we would like to clarify that our correlation analyses do not prove that empathy can be predicted at the level of the individual. We can consider the discussion on the prediction of empathy at the level of the individual only as speculative ideas. Confirmation of whether or not we can successfully predict empathy from online social network structure will require additional work, such as by evaluating the performance of a binary classification system (classifying individuals as empathic/non-empathic) through ROC curves (Zweig and Campbell, 1993). The contributions of this work, and our primary interest, are in the insights at the level of the population that the study reveals.

### *A note on the mediation analysis*

In section IV *C* of the paper, we reported a mediation analysis to test whether empathy acts as a mediator in the relationship between network structure and social capital. We then suggested that, while the analysis supported the hypothesis that empathy plays a mediation role, careful confirmation of the direction of causality will require longitudinal assessment. Here we elaborate on this issue. First, we note that mediation analysis by itself cannot prove the direction of a causal relationship, but can lend credibility to our hypothesis of how these variables are related, which must be founded on prior work or some form of valid reasoning. We reasoned that network structure is likely to affect empathy, since interaction with diverse individuals affects empathy (Boisjoly et al., 2006). As empathy also affects social capital, it is possible that part of

the translation of network structure into social capital happens through the mediation of empathy.

However, as we have noted in section II of the paper, there is also evidence to suggest that empathy might affect network structure, since empathic individuals are more comfortable with individuals of different backgrounds (Galinsky, 2002). As a result, it is also possible that the relationship between empathy and social capital might be partly explained by the relationship between empathy and network structure, establishing network structure as a mediator in the relationship between empathy and social capital. To fully understand the feedback effects between empathy and network structure would require a longitudinal assessment of individuals' empathic ability as their network structures evolve over time. Nevertheless, regardless of the magnitude of causality in either direction between empathy and network structure, these findings suggest that empathy, network structure and social capital to an extent vary together.

#### *Future Work and Practical Implications*

Since online social networks play an important role in the creation of social capital, examining the relationship between fundamental aspects of humans' behaviour and social capital in the context of online social networks, as we have done in the previous chapter and in this chapter, can shed valuable insights for the design of these systems. While these chapters are initial steps in this direction and we have not focussed directly on practical applications at this stage, here we discuss briefly the implications that this line of work can have by considering one particular application of online social networks : automated (or semi-automated) friend suggestions.

When attempting to assist a user expand his or her social network in an online social network like Facebook, it is important that we ask why might we want to do this. We argue that the primary objective of such applications should be to effect social capital, and more generally the well-being, of the individual or the community. By setting well-being as our guiding criterion, we can begin to understand more clearly the value and limitations of automated or semi-automated friend suggestion applications, and begin to explore how to affect well-being through friend suggestions. While the problem of "link-prediction" is to predict which individuals might already have some degree of acquaintance with each other in the offline (eg. Cranshaw et al, 2010), such as offline friends or "familiar strangers" (Paulos and Goodman, 2004), the goal we are referring to

is to understand which links when activated might have a positive impact on the well-being of the individuals.

Prior work in the social sciences such as Burt's theory of structural holes (Burt, 1995) give us some pointers in this regard by suggesting a relationship between social capital and features of network structure such as structural holes, which the analyses from our studies support in the context of online social networks. However, a more nuanced understanding in terms of how individuals' behavioural tendencies affect the relationship between network structure and social capital would be more useful towards understanding friend suggestion applications. We have attempted to explore this in terms differences between individuals in personality traits. This is only a first step in this direction, and we need to examine more thoroughly which universal behavioural tendencies might be most impactful in the effect they have on the relationship between network structure and well-being. For example, childhood attachment style, which reflects the level of security experienced in the bond with parents in early childhood, can affect later capacity to make affectional bonds (Bowlby, 1977). Thus, it is quite plausible that a particular strategy for expanding one's network might be helpful for an individual with secure attachment (which can be considered a healthy style of relating) but inappropriate for an individual with disorganised attachment (which can be considered a dysfunctional style of relating). Since different attachment styles occur with different frequencies in the population (Benoit, 2004), the trends we find in the relationship between network structure and social capital or well-being can be a reflection of that bias.

The point we wish to make here is that while there can indeed be some broad trends in the relationship between network structure and social capital such as the effect of structural holes, this might conceal the different ways in which different individuals might be able to have their needs for well-being met based on the universal behavioural tendencies that operate in them. Thus, studying the effect of universal behavioral tendencies on the relationship between social network structure and social capital relationship can help us understand the value and limitations of adopting a common friend suggestion strategy for all individuals in the social network, and provide insights on how to improve these applications.

Similarly, closely examining the role of empathy in relation to friending and social capital needs can help us understand under which circumstances and to what extent friend suggestion applications can be helpful. For example, it is possible that individuals with high empathy might benefit in terms of social capital from an expansion in their social network, but for individuals low on empathy, an increase in the number of friends might not have the same positive effect, or might even be detrimental to their well being if the added friends are of a different background. This might be examined through a longitudinal study. Without such a closer examination, we can only base our friend recommendation applications on the assumption that expanding one's network is likely to be helpful in general. Both from a scientific and a practical perspective, it is important that we understand the degree to which this assumption might fail to hold.

#### *Conclusion*

In this chapter we set out to study the relationship between empathic ability and online social network structure, and the role of empathy in the relationship between online social network structure and social capital reported in the previous chapter. The results suggest that empathy is related to online social network structure, as we had conjectured based on prior literature. In particular, structural holes and the number of friends were associated with cognitive empathy. Further, our analysis revealed a mediation pathway between structural holes, empathy and bridging social capital, suggesting that these variables to some extent vary together. Overall, the study suggests that empathy is to an extent reflected in network structure, and to that extent social networks analysis can be used as a lens through which we can study this human ability that is fundamental to successful relationships and social capital.

### Chapter References

The following are the references cited in this chapter outside of the embedded paper.

Benoit, D. (2004). Infant-parent attachment: Definition, types, antecedents, measurement and outcome. Paediatrics & child health, 9(8), 541.

Boisjoly, J., Duncan, G. J., Kremer, M., Levy, D. M., & Eccles, J. (2006). Empathy or antipathy? The impact of diversity. The American economic review, 96(5), 1890-1905.

Bowlby, J. (1977). The making and breaking of affectional bonds: I. Aetiology and psychopathology in light of attachment theory. Br J Psychiatry. 130:201–10.

Burt, R. S. (1995). Structural Holes: The Social Structure of Competition. Harvard University Press.

Cranshaw, J., Toch, E., Hong, J., Kittur, A., & Sadeh, N. (2010). Bridging the gap between physical location and online social networks. In Proceedings of the 12th ACM international conference on Ubiquitous computing (pp. 119-128).

Galinsky, A. D. (2002). Creating and reducing intergroup conflict: The role of perspective-taking in affecting out-group evaluations. In M. A. Neale, E. A. Mannix, & H. Sondak (Eds.), Toward a phenomenology of groups and group membership (Vol. 4, pp. 85–113). Greenwich, CT: JAI.

Paulos, E. and Goodman, E. (2004). The familiar stranger: Anxiety, comfort, and play in public places. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'04). ACM, 223–230.

Preece, J. (1999). Empathic communities: Balancing emotional and factual communication. Interacting with computers, 12(1), 63-77.

Wölfer, R., Cortina, K.S. & Baumert, J. (2012). Embeddedness and empathy: How the social network shapes adolescents' social understanding. Journal of Adolescence 35, 1295–1305.

Zweig, M. H. and Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clinical Chemistry 39 (8): 561–577.

# 7. Conclusion

We began the thesis by outlining the role of universal behavioural tendencies in online interaction behaviours, and the effect that these two factors in turn have on social capital. This discussion culminated in our overarching research question for the thesis, which served as the basic motivation for us to present the work that we conducted in the subsequent chapters with the purpose of addressing different aspects of this overarching research question. In this chapter we return to our overarching research question and discuss in what ways we have managed to address it, both in terms of the insights we have gathered towards understanding the interplay between universal behavioural tendencies, online interaction behaviour and social capital, and in terms of the methodologies we have explored that can be used to take this investigation further (Section 7.1). We also discuss the broader limitations of our work, and provide some pointers on the directions that future work can explore to tackle those limitations (Section 7.2).

## 7.1 Contributions

Here we discuss in what ways our work has addressed the overarching research question for this thesis:

*How do universal behavioural tendencies influence online interaction behaviour, and how do they both in turn affect social capital?*

The two specific proxies for online interaction behaviour that we consider in this thesis are disclosure behaviour and social network structure.

#### *Online Disclosure and the Tendency to Reciprocate*

Our first study (Chapter 3) examined the effect of one important universal behavioural tendency, the tendency to reciprocate, on disclosure behaviour in an online social network. Specifically, we examined the effect of the tendency to reciprocate on individuals' willingness to share personally identifiable information with strangers, in an online social network in which interactions are typically between strangers. Our study

resulted in two main findings that contribute towards the understanding of the role of reciprocity in online disclosure.

The first finding is that individuals do indeed reciprocate the disclosure of personally identifiable information with strangers in online social networks. While we conjectured this result based on prior work on self-disclosure in face to face interactions, we feel that this empirical verification of the conjecture in the context of online interactions is valuable. The implication of this finding is that the tendency to reciprocate nudges users in online social networks to share their information with other users when those users share their information. While ours was a controlled study on one particular online social network, Livemocha, the advantage of working with a universal behavioural tendency such as reciprocity is that it is likely to operate across a wide range of circumstances since it is fundamental to human nature. Therefore, the tendency to reciprocate is likely to provide a nudge towards sharing one's information across numerous online social network contexts, whether or not that nudge is strong enough to result in a disclosure in any particular situation in an online social network, given the contextual forces operating within that situation.

The second finding is that the tendency to reciprocate is triggered not only when the initial disclosure is shared in a one-to-one and directed manner, but also when the initial disclosure is broadcast and directed to no one in particular. The implication of this finding is that the tendency to reciprocate can kick in even when information is disclosed to a group or community of people as a whole and then an explicit request is made to an individual from this group. This is a surprising result that we might not have expected based on the understanding of prior literature in the context of face to face interactions. Given that in numerous online social networks, such as Facebook, a significant proportion of information sharing is done through broadcast channels in which the information is directed to nobody in particular, this finding contributes towards our understanding disclosures in a wide range of online social networks.

The findings also have implications for privacy. This work demonstrates the vulnerability of users against attempts to trick them into revealing information by exploiting this social norm. Inferring or linking personal information such as that obtained in the current study would typically be an important first step in a malicious party's attempt to exploit a user. The awareness and understanding of the phenomenon

of self-disclosure resulting from reciprocity gives us the perspective that enables us to look for solutions to the problem (in this case, a malicious party exploiting a behavioural tendency to elicit information) through appropriate means, such as by educating users or by appropriately priming them to exercise caution with certain users (for example, by making the reputation score more salient) as suggested in Chapter 3.

***Studying the Interplay between Disclosure Patterns and Network Structure***

In our next study (Chapter 4) we jointly examined the two proxies for interaction behaviour that are central to the focus of this thesis, disclosure patterns and network structure. We articulated hypotheses to be tested on the data we extracted for these two proxies of interaction behaviours, partly relying on ideas based on universal behavioural tendencies such as homophily and reciprocity, and partly relying on other observations. From the perspective of our thesis, the key finding of this work is that disclosure patterns and network structure are related - users who are more central to their social network are more likely to share information about their location, and hence are more "vocal". In addition, the findings show that given a target, that target's friend who either has many friends or many common friends with the target is more likely to be trusted by the target. Our analysis also revealed patterns in the dynamics by which triads of friends share their location with each other. These findings can be useful both in understanding how privacy attacks can be engineered and how they can be prevented, and therefore one of the practical implications is that the findings can inform the development of automated protection and suggestion mechanisms that will make the sharing of real-time location safer and more useful.

Beyond the findings of how disclosure patterns and network structure are related in the context of the particular social network we examined, Locaccino, our work contributes towards establishing a general method that can be replicated in other systems. In chapter 4, we explained why the combined examination of the two proxies of disclosure patterns and network structure for online interaction behaviour is little explored - it is difficult to manually capture and measure instances of self-disclosure between individuals on a large scale. By using the privacy policy (or privacy settings) of users in the online social networking system, we were able to automatically generate an "openness" metric, that reflects disclosure behaviour for users across the system. This simple idea enables us to study disclosure at the much larger scales at which we are

typically able to study network structure from the graphs of online social networks today.

Therefore, this study demonstrates the potential of our methodology for the combined analyses of disclosure behaviour and social network structure on a large scale across different online social networking systems. The method we have described can be provide rich insights into the interaction patterns at the population scale when the social networking system provides mechanisms that enables the user to control, for each of her friends, how much of the information that she shares is visible to that friend. It is increasingly the case in today's online social networking systems that they incorporate such control mechanisms as they evolve. For example, both Facebook and Google Plus provide mechanisms that attempt to make it convenient for users to pick out a subset of friends to share a particular post or piece of information. Thus the method we suggest to study disclosure information and network structure in conjunction can be potentially applied across a number of social networking systems.

#### *Network Structure, Personality and Social Capital*

Our third study (Chapter 5) examined the relationship between online network structure and social capital. Our motivation for this work was to touch upon two related issues regarding the benefits one receives from his or her social ties. The first is a body of work in traditional social science which suggests that the structure of the social ties around individuals affects the benefits they receive on account of those social ties (eg. Granovetter, 1973; Burt, 1995). In particular, according to structural holes theory (Burt, 1995), structural holes should be positively related to bridging social capital. We were interested in examining the veracity of this hypothesis in the context of online social networks. Our finding is that online network structure, based on data we collected from the Facebook social network, is indeed related to social capital in the way the theory of structural holes predicts: the extent of structural holes in our participants' online networks were positively related to bridging social capital. In addition, the number of isolated friends was significantly related to bridging social capital. Overall, the evidence for the relationship between network structure and social capital in our analysis was stronger for bridging social capital than for bonding social capital.

The second issue we investigate is whether individual differences affect the relationship between network structure and social capital, and how. While numerous kinds of individual differences such as communication skills can affect the relationship between network structure and social capital, we chose to examine the effect of personality traits, as characterized by the big five model of personality (Costa and McCrae, 1992), as our interest was focussed on exploring the effect of more fundamental universal behavioural tendencies on this relationship. Our results provide a number of provisional insights on effect of personality traits on the relationship between network structure and social capital. For example, introverts benefit more strongly than extroverts in terms of bonding social capital from networks with higher transitivity. This is an intuitive finding which suggests that tightly knit clusters of ties benefit introverts more strongly than extroverts for bonding needs. On the other hand, our data also suggested relationships between variables that were not quite intuitive. For example, those who were low in agreeableness benefitted more in terms of bridging social capital from having isolated friends than those who were high in agreeableness. Overall, our results on the role of personality in the relationship between network structure and social capital provide a number of interesting provisional insights that we can use to direct subsequent explorations and hypotheses in future work.

We also explain the context of the current activity within the research community in which we see the value of this direction of work of exploring the use of social networks analysis to understand the relationship between online social networks and social capital. Since social networks analysis has a long history in the social sciences, and there is a body of work on network structure grounded in theories of how individuals and groups interact and affect each other (e.g., Berkowitz, 1982; Granovetter, 1983), it lends itself naturally to the validation of these theories of human interaction. This makes it valuable to help us make sense of the underlying mechanisms that govern these interactions. Therefore, this direction of investigation can ultimately contribute towards an integrated understanding of these phenomena based on theoretical foundations that is arguably lacking in the current literature in HCI on the relationship between online social networks and social capital.

***Understanding Empathy Through Network Structure***

Our final study (Chapter 6) examined the relationship between empathy and network structure, and the role that empathy plays in the relationship between social capital and network structure. Empathy is an important human ability that allows us to understand the other person's perspective and emotions. It facilitates meaningful interactions between individuals and is therefore fundamental to successful human relationships. Since empathy is closely related to human relationships, we hypothesized that empathy might be reflected in social network structure, and consequently our goal for this work was to explore the use of social networks analysis in understanding empathy in the context of online social networks. Our results based on data from participants on the Facebook social network supported our hypotheses on the relationship between empathy and network structure. In particular, both structural holes and the number of friends of individuals were positively related to cognitive empathy. This lends support to the idea that social network structure can be helpful in understanding empathy in the context of online social networks.

Extending our examination of social capital from the previous study (Chapter 5), we also sought to understand the role that empathy plays in the relationship between network structure and social capital. First, our findings suggest that empathy is positively related to social capital. In particular, cognitive empathy showed a positive relationship with both bridging and bonding social capital. Further, emotional reactivity, which can be considered an emotional component of empathy, showed a significant relationship with bonding social capital.

Further analysis revealed a mediation pathway between structural holes, cognitive empathy and bridging social capital. While ascertaining the extent of the causality in each direction between network structure and cognitive empathy will require further study, our results suggest that these variables (structural holes, cognitive empathy and bridging social capital) to an extent vary together. This is an important result in our quest to understand how social capital is accrued, as it suggests that network structure and empathy play a joint and interconnected role in the creation of social capital.

## 7.2 Limitations and Future Work

We have discussed in detail the limitations of each study within their respective chapters in the thesis. Here we touch upon some of the broader limitations of the work presented in this thesis, and the directions that future work can explore to address them.

First, we note that we chose to study reciprocity in the study presented in Chapter 3 as it is an important universal behavioural tendency in terms of the effect it has on people's interactions, and we arrived at specific findings of how reciprocity affects disclosure behaviour. We could similarly study other universal behavioural tendencies and how each one affects online interaction behaviour, and such studies can enhance our understanding of the problem space. However, merely following such an approach might lead only to fragmented insights into these phenomena, and the value of such work might be greatly enhanced if they can be integrated and seen from the perspective of a bigger picture of how they might relate or contrast with each other. For example, it may be possible that different behavioural tendencies result in similar behaviours, but are triggered under different conditions, as might be the case with reciprocity and imitation (Venkatanathan and Kostakos, 2011). It would be valuable to develop a larger framework or theory (or theories, as there can be multiple ways of looking at the bigger picture) that incorporates behavioural tendencies, elements of context, social network behaviour and social capital. For example, a classification of behavioural tendencies and a classification of contextual factors and the effects that different classes of behavioural tendencies, coupled with different contextual factors, have on social network behaviour and social capital can allow us to explore this space in a systematic manner and also allow us to understand these factors from the perspective of a larger picture. Thus it would be fruitful for future work to study these phenomena while also laying an emphasis on how they compare and fit together into a larger framework or theory.

One of the main limitations of the work in chapters 4, 5 and 6 based on social networks analysis is that to an extent they simplify individuals and the ties between them into nodes and links. This simplification masks out a range of finer qualitative details such as the nature of close and intimate relationships. Nevertheless, this simplification is valuable as it allows us to identify and study patterns of behaviour at the level of the population. Moreover, we have enriched our network representations to some extent based on qualitative information, such as with disclosure information (Chapter 4) and

with personality and empathic ability (Chapters 5 and 6). Thus, we do not reduce humans to plain nodes on a graph, but rather depict them, and the ties between them, as nodes and links with various properties.

One other limitation of the method we have suggested for the combined analysis disclosure patterns and network structure (Chapter 4) is that of the availability of privacy policy data to the average researcher. Since this data can be considered particularly privacy sensitive, complete access to such information might only be available to companies and developers of social networking sites. Thus, to some extent, the possibility of carrying such an investigation might be restricted to the developers of these social networking sites. However, this difficulty might be overcome by accessing the history of posts shared by users, and to which of their friends these posts were shared. This information might be used to develop an openness metric, based on which the subsequent analyses that we have carried out can be performed. Some online social networking sites such as Facebook provide mechanisms that enable the user to provide access to this data to third party application developers, which can be utilized by researchers. Further, other kinds of tie information can also be embedded into the links during the analysis of the network, such as measures of tie strength inferred from various kinds of communication and interaction between users along the lines of the work of Gilbert and Karahalios (2009). This kind of data can also be coupled with our work on ego-centric networks (Chapters 5 and 6). For example, tie strength can be used as edge weights in the calculation of betweenness centrality, or can be used in place of the "time" that the ego spends with his contacts in the calculation Burt's constraint metric. Thus coupling social network structure with various properties embedded on to the links can enable us to perform more nuanced analyses.

Finally, in our treatment of the various concepts we encountered in chapters 5 and 6, we only considered pairwise correlational analyses, and sometimes examined a third variable in the role of a mediator or moderator in the relationship between two variables. While this has been useful to draw a number of insights based on the various network metrics we calculated and the questionnaire measures we collected, clearly the variables we are studying simultaneously affect each other in various ways. Therefore, future work building on this should attempt to analyze these variables simultaneously, so as to understand more fully how these variables affect each other. For example, in order to carry out the theoretical modeling and simultaneous analysis of multiple

factors, we can use methods such as Bayesian networks (Ben-Gal, 2007), or structural equation modeling (Bollen, 1998) as our group has started to explore (Liu et al., 2014). We suggest that by using these advanced statistical tools along with the methods we have explored in this thesis, there is much potential for HCI research towards further understanding universal behavioural tendencies and social capital in the context of online social networks.

**Chapter References**

Ben-Gal, I. (2007). Bayesian networks. Encyclopedia of statistics in quality and reliability. John Wiley & Sons Ltd.

Berkowitz, S. D. (1982). An introduction to structural analysis: The network approach to social research. Toronto: Butterworth.

Bollen, K. A. (1998). Structural equation models. John Wiley & Sons, Ltd.

Gilbert, E., & Karahalios, K. (2009). Predicting tie strength with social media. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 211-220). ACM.

Granovetter, M. (1983). The strength of weak ties: A network theory revisited. Sociological Theory 1 (1983), 201–233.

Liu, Y., Venkatanathan, J., Goncalves, J., Karapanos, E., & Kostakos, V. (2014). Modeling What Friendship Patterns on Facebook Reveal About Personality and Social Capital. ACM Transactions on Computer-Human Interaction (TOCHI), 21(3), 17.

Venkatanathan, J. and Kostakos, V. (2011). Privacy in a networked world: effects of reciprocity and imitation on location sharing. Proc. CHI workshop on Networked Privacy, Vancouver, Canada.