Tampere University

JAYESH PRAKASH GUPTA

Identification & Role of Implicit Social Ties from Social Data

Tampere University Dissertations 722

Tampere University Dissertations 722

JAYESH PRAKASH GUPTA

Identification & Role of Implicit Social Ties from Social Data

ACADEMIC DISSERTATION
To be presented, with the permission of
the Faculty of Management and Business
of Tampere University,
for public discussion in the Auditorium S2 (SA203)
of the Sähkötalo, Korkeakoulunkatu 3, Tampere,
on 16 December 2022, at 13 o’clock.

ACADEMIC DISSERTATION
Tampere University, Faculty of Management and Business
Finland

Responsible supervisor and Custos	Professor Hannu Kärkkäinen Tampere University Finland	
Pre-examiners	Professor Derek L. Hansen Brigham Young University United States	Associate Professor Kaisa Henttonen University of Eastern Finland Finland
Opponent	Ph.D. Kaisa Still University of Oulu Finland	

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

Cover design: Roihu Inc.

ISBN 978-952-03-2695-1 (print)
ISBN 978-952-03-2696-8 (pdf)
ISSN 2489-9860 (print)
ISSN 2490-0028 (pdf)
http://urn.fi/URN:ISBN:978-952-03-2696-8

Carbon dioxide emissions from printing Tampere University dissertations have been compensated.

PunaMusta Oy – Yliopistopaino
Joensuu 2022

karmany-evādhikāras te
mā phaleshu kadāchana
mā karma-phala-hetur bhūr
mā te sago 'stvakarmai

You have a right to perform your prescribed duties, but you are not entitled to the fruits of your actions. Never consider yourself the cause of the results of your activities, and never be attached to not doing your duty.

Bhagavad Gita Chapter 2, Verse 47

ACKNOWLEDGEMENTS

This research was conducted between 2017 and 2022, first in Tampere University of Technology and later in Tampere University under the Doctoral Programme in Business and Technology Management in the Unit of Information and Knowledge Management. During this time, I have been financially supported by the Doctoral Program of Tampere University, Finnish Cultural Foundation, and TES, whom I want to warmly acknowledge.

The initial spark for this dissertation came from my supervisor Professor Hannu Kärkkäinen, who, after supervising my master's theses, kindly offered me an opportunity to continue towards a doctoral degree. I would thus like to express my gratitude to Professor Hannu for introducing me to the academic world and its community. Your supervision and advice have been invaluable. I would thank you from bottom of my heart for having faith in me and keeping the Ph.D. ship afloat during turbulent times.

I would also like to thank Associate Professor Raghava Rao Mukkamala (Copenhagen Business School, Denmark) for always being a guiding star and making me believe that there is light at the end of the Ph.D. tunnel. I am also grateful to Associate Professor Hongxiu Li (Tampere University, Finland), for helping me transform vague ideas into real academic papers.

I also wish to express my gratitude to my pre-examiners, Professor Derek L. Hansen from Birgham Young University, United States, and Associate Professor Kaisa Henttonen from University of Eastern Finland, Finland, for their insightful comments and suggestions related to this dissertation, as well as Dr. Kaisa Still from the University of Oulu for acting as the opponent in the upcoming defense.

I would like to thank Professor Ravi Vatrapu (Ryerson University, Canada) for not only being a good collaborator but also a great host during my research exchanges at Center for Business Data Analytics, Copenhagen Business School, Denmark. I would like to thank all my collaborators - Professor Thomas Olsson (Tampere University, Finland), Associate Professor Henri Pirkkalainen (Tampere University, Finland), Associate Professor Abid Hussain (Copenhagen Business School, Denmark), Dr. Jari Jussila (HAMK, Finland) and Osku Torro. A special thanks to Dr. Jukka Huhtamäki (Tampere University, Finland) for introducing me to the world of social network analysis. A big thanks to Dr. Karan Menon for not only being a great office mate but also a person from whom I could always seek advice and help. During my doctoral journey I was also fortunate to work in projects like COBWEB (Academy of Finland) and Big Match (Business Finland). These projects gave me chance to closely collaborate with great colleagues like Kate (Ekaterina Olshannikova), Erjon, Yao Peng and other wonderful project mates. I would like to thank you all for being part of this journey.

I would like to thank Professor Nina Helander and Professor Samuli Pekkola to support my work as unit heads. I also had the pleasure of working with some amazing colleagues at the university. I would like to thank Pasi, Ilona, Maija, Prasanth, Dang, Sameer, Prasanna, Mikko, Heidi, Kestutis, Margarita, Professor Juho Kanniainen for some memorable outings, christmas parties, and coffee/ cake sessions. I have also been lucky to have friends outside of work around the world. I would like to thank Praveen, Rohan, Pritha, Madan, Aditya, Vidisha, Ramya, Hamza, Ram Krishna, Manu, Anant, Kamal, Kabira, Nishit, and all my floorball friends. Many people crossed my path during my PhD journey that led me here I can not mention everyone, but I am thankful to all of you.

I have also been lucky to have a caring family. I would like to thank my parents and my brother for all their support. I would also like to thank my in-laws (my wife's parents and her brother) for having trust in my abilities and encouraging me to move forward.

Finally, I would like to thank my wife Shruti without whom this endeavor would not have been possible. I know it has not been an easy journey for you. Thank you for being my pillar of strength, always encouraging and keeping me sane. I am blessed to have such a wonderful person in my life. Thank you for being a constant support

Tampere 17 Nov 2022

Jayesh Prakash Gupta

ABSTRACT

The concept of social ties was introduced by Granovetter through the seminal paper titled" Strength of weak ties". Across the past five decades, this topic has attracted much attention from both academics and practitioners. In the past two decades, the rapid increase in digitization and new modes of communication have led to collecting and analyzing data about people. One of the most popular sources for such large and granular data about people is social media platforms. The rise in the popularity of social media in the past 15 years has resulted in many research studies that have used social media data to understand a lot of different phenomena. Some of this research has focused on using social data, including social media data, on identifying different kinds of social ties online and the role these social ties play in various contexts. Over the past decade, many different approaches and models have been built to identify social ties using social media data. These methods have been built using private data and explicit social relationship data of users' social media platforms. However, in the past few years, it has become nearly impossible to access this kind of social media data due to the changes in the business models of the social media platforms and the introduction of new privacy laws like GDPR.

This thesis aims to identify the social ties from publicly available social data and study the role of the identified social ties in different contexts like business conferences and business phenomena. In order to achieve this research objective, three separate studies were conducted. The first two studies were single-case case studies, while the third was an experiment where two different sets of hypotheses were tested using empirical data. All three studies used publicly available social media data related to a specific context. The first study used a large dataset related to a game developer community on Facebook. The second study used social media data related to a business event from Twitter and Facebook. The third study used a large dataset associated with social media data about crowdfunding projects from Twitter.

This study adds to the existing literature related to identifying social ties from social media data in multiple manners. The thesis illustrates a novel approach based on reciprocal interaction for filtering relevant social ties from large publicly available social media data. The thesis also contributes to the understanding of the role multiple social media platforms play in an event. Thus, showing the impact this can have on identifying social ties from publicly available social media data in case of an event. The dissertation adds to the existing literature about the role social ties have towards crowdfunding success. The thesis shows that implicit social ties, in general, positively impact crowdfunding project success. In addition, the thesis has practical implications for designers of conference recommendation systems. The dissertation also has implications for the crowdfunding project owners and the crowdfunding project campaign designers.

CONTENTS

1 INTRODUCTION

1.1 Background and rationale

The concept of social ties was introduced by Granovetter in his seminal work "The Strength of Weak Ties". In this study, he defined the concept of tie strength and based on it defined two kinds of social ties - strong ties and weak ties. (Granovetter 1973) Over the past five decades, this topic has attracted a lot of attention from both academics and practitioners. The concept of tie strength and the related social ties has been now been used in a variety of fields including social science, economics, computer science, information science, innovation management, business, and many other areas. The role of social ties has been used to analyze a vast variety of phenomena ranging from the original context of job search (Granovetter 1973); and now includes many other contexts like knowledge transfer (Levin and Cross 2004), content sharing (Zhan Shi, Huaxia Rui and Whinston 2014), information diffusion (Yi, Z. Zhang and Gan 2018), social influence (Bakshy et al. 2012), large scale networked experiments (Aral and Walker 2014), information propagation (Zhao et al. 2012), collaborations (Dahlander and McFarland 2013), crowdfunding (Borst, Moser and Ferguson 2018) and other different contexts. Thus, the concept of social ties has provided an important perspective in understanding and explaining a variety of phenomena.

The rapid increase in digitization and adoption of new information and communication technologies by people has enabled businesses and researchers to collect and analyze data about individuals at a scale which was unthinkable earlier (Davis et al. 2016). Over the past decade, the rise in social data especially social media data has provided researchers with a new data source for the identification of social ties. This has resulted in a lot of studies that have used data from different social media platforms to identify the different kinds of social ties using this data. One of the earliest social ties identification model based on social media data was developed by Gilbert and Karahalios 2009. This model used more than 70 measures based on Facebook, to identify the strong and weak social ties of different Facebook users. Another study developed a logistic regression model based on Facebook data to identify different kinds of social ties (Jones et al. 2013). Some studies used supervised machine learning methods (Kahanda and Neville 2009) while others used unsupervised machine learning methods (Xiang, Neville and Rogati 2010) for the identification of different kinds of social ties. Some of the most important social ties identification models based on social media are also provided in the study by Gupta, Kärkkäinen, Torro et al. 2019.

An important similarity in all the above studies and most of the other studies related to the social ties identification models based on social media data is their reliance on explicit relationship data about the social media platform users. However, in the past few years, the introduction of new regulations (like GDPR) and also the increasing restrictive data access policies of the social media platform (Hogan 2018) have made it almost impossible to access the explicit relationship data of the users of the social media platforms (Gupta, Kärkkäinen, Torro et al. 2019). Thus, these changes have made it impossible to directly use the existing social ties identification models based on social media data.

The above mentioned challenges in the use of social media data for the identification of social ties have resulted in a need to look for other ways to identify social ties. The social networks based on social media data can be broadly classified into two types - explicit social networks and implicit social networks (Zhou, Duan and Piramuthu 2014). The explicit social networks are based on the explicit relationship data available from the social media platforms while the implicit social networks can be based on any kind of social media data except the explicit relationship data. Even with the current restrictive data access policies of social media platforms, it is still possible to access a lot of publicly available social media data related to user interaction, conversation, and other textual data. This kind of publicly available social media data provides an opportunity to create different kinds of implicit social networks and use these implicit social networks for the identification of different kinds of social ties.

In the past, some studies have created implicit social networks and used them for different purposes. For example, a few studies have used the implicit social networks for understanding the co-learning and informal learning behavior of conference par-

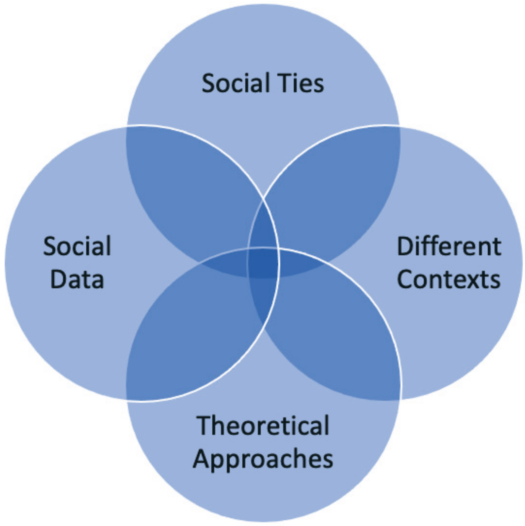

Figure 1.1 Targeted contribution of the dissertation

ticipants (Aramo-Immonen, Jussila and Huhtamäki 2015; Aramo-Immonen, Kärkkäinen et al. 2016). At the same time there are different theoretical approaches based on the network structure and the position of the different nodes in the network (like structural holes theory) which can be used to develop different measures and methods to identify social ties. However, there is very little research which has focused on using implicit social networks to identify different kinds of social ties.

1.2 Research aim and focus area of the thesis

The aim of this thesis is to identify the social ties from social data and also to study the role of implicit social ties in different contexts like business conferences and business phenomena.

In summary, the contribution of this thesis concerns the interaction of social ties, social data, different theoretical approaches related to the identification of social ties, and also the different contexts in which these social ties can be identified and used. There is existing research on social ties and how they can be identified from the social data (Fogués et al. 2013; Gilbert 2012; Gilbert and Karahalios 2009; Jones et al. 2013; Kahanda and Neville 2009; L. Fogues et al. 2018; Liberatore and Quijano-Sanchez 2017; Xiang, Neville and Rogati 2010). There also exists research which has used different theoretical approaches and perspectives for identification of different kinds of social ties from social data (Burt 1992; De Meo et al. 2014; Gupte and Eliassi-Rad 2012; Haythornthwaite 2002; Huang et al. 2015; Krackhardt 1992; Levin, Walter and Murnighan 2011; Marsden and Campbell 2012; Mattie et al. 2018; Onnela et al. 2007; Petroczi, Bazsó and Nepusz 2007; Zhan Shi, Huaxia Rui and Whinston 2014). At the same time, there is existing research which has studied the role of social ties in different contexts in order to understand various phenomena (Aral and Walker 2014; Borst, Moser and Ferguson 2018; Hong, Hu and Burtch 2018; Kang, Jiang and C. H. Tan 2017; Levin and Cross 2004; Mcguire and Bielby 2016). Hence, this dissertation focuses on the intersection of all these topics mentioned above and shown in Figure 1.1

On the basis of the above research gap description this dissertation addresses the following three research questions:

RQ1: How can different kinds of implicit social ties be identified using the publicly available social data and big social data ?

RQ2: How can the potentially useful implicit social ties be identified using social data and big social data in different knowledge work context like professional conferences, research conferences and other similar contexts?

RQ3: How can the potentially useful implicit social ties, identified from existing social data and big social data, be used in different professional contexts like business decision-making or business phenomena?

1.3 Original Research Plan

Originally this thesis was planned to be an article-based dissertation. However, due to some unforeseen reasons, it was decided to convert the dissertation into a monograph. This section presents the list of publications that were planned to be part of the article-based dissertation. Table 1.1 below provides the titles of the different articles, the publication outlets of the articles, and the research questions that these articles were intended to answer. The structure of the thesis in the current form of a monograph is presented in the next section.

Table 1.1 Original research plan for the dissertation

Paper	Outlet	Research Question
Identifying weak ties from publicly available social media data in an event	Mindtrek 2016 (Conference)	RQ 2
Integrating micro-level interactions with social network analysis in tie strength research: the edge-centered approach	Mindtrek 2017 (Conference)	RQ 1
Identifying different types of social ties in events from publicly available social media data	KMIS 2019 (Conference)	RQ 2
Revisiting Social Media Tie Strength in the Era of Data Access Restrictions	KMIS 2019 (Conference)	RQ 1
Role of Implicit Social Ties in Crowdfunding Success	Under Review (Journal)	RQ 3
Social Ties' Role in Crowdfunding Projects' Success: A Project-Owner and Project-Backer Perspective	Under Review (Journal)	RQ 3

1.4 Structure of the thesis

To meet the objectives of this research (see, 1.2, three different studies were conducted. The first two studies were single-case case studies. In contrast, the third study was an empirical experiment based on testing specific hypotheses. The third study was conducted in two parts and tested two different sets of hypotheses. The first two case studies were selected and conducted following the guideline suggested by R. K. Yin 2018. The third study was an experiment and was designed keeping in view the principles and guidelines suggested by Saunders and Thornhill 2019.

The thesis is comprised of six chapters. Chapter 1 explains the background and rationale, research aim and intended contributions of the thesis, the thesis structure, and finally, the original research plan for carrying out the thesis. Chapter 2 presents the key concepts related to the research topic and provides an overview of the current understanding of the research topic. The final section provides the reason for selecting the different research contexts. Also, it provides an overview and the research gap about the selected research contexts.

Chapter 3 describes how the research was conducted in practice and presents the selected research strategy's methodological choices. The chapter also introduces the different studies conducted as a part of this dissertation. The chapter provides details about the different data sources and analysis methods used in conducting these studies. The final section of this chapter provides a brief summary of the different research strategies, methods, and datasets used in this dissertation. Chapter 4 forms the empirical part of the study and describes the results of the research. The results of all three studies are presented in the different sections of this chapter.

Chapter 5 discusses the findings of the different studies based on the results of the individual studies. Chapter 6 concludes the thesis. First, the conclusions and academic contributions of the research are summarized. Second, the managerial implications of the dissertation are presented. Third, the evaluation of the study is presented. Finally, the limitations of the study and the future research opportunities are presented.

2 THEORETICAL BACKGROUND

2.1 Theories related to social ties

2.1.1 Concept of tie strength and different types of social ties

The concept of tie strength was introduced by Mark Granovetter in his seminal paper titled "Strength of weak ties". According to Granovetter 1973, the tie strength can be defined as "a (probably linear) combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie". In the original definition provided by Granovetter, the tie strength was used to understand the different interpersonal relationships. In other words, the concept of tie strength provided the degree of closeness between two individuals (Gupta, Menon et al. 2016). However, Granovetter left the more precise definition of tie strength to future work.

In his original study, Granovetter used the concept of tie strength to understand how people looking for a new job found this information from there social network (Granovetter 1973). Over the decades, the concept of tie strength has been used to study various social phenomena beyond the original job-seeking context. (e.g. innovation and creativity (Sosa 2011), knowledge transfer (Levin and Cross 2004), information diffusion (Gilbert and Karahalios 2009), and content sharing (Aral and Walker 2014; Zhan Shi, Huaxia Rui and Whinston 2014). At the same time the tie strength measurement has been extended from its original use at the interpersonal level to organizational and inter-organizational levels as well (e.g. Granovetter 1973; Wegge et al. 2015).Over the years, different studies have used different measures to either measure tie strength or have used different measures as a substitute for tie strength. Some of the most widely used offline (non-social media based) measures are presented in table 2.1

Table 2.1 Some common measures of tie strength

Measure	Related Studies
Closeness	Blumstein and Kollock 1988; Daly and Haahr 2009 **Echebarria2013LimitsRegions**; Levin and Cross 2004 Marsden and Campbell 1984; Marsden and Campbell 2012 Mathews et al. 1998; Mcguire and Bielby 2016 Nitzan and Libai 2011; Verlegh et al. 2013
Frequency	Dahlander and McFarland 2013; Granovetter 1973 N. Lin, Dayton and Greenwald 1978; Mcguire and Bielby 2016 Nitzan and Libai 2011; Onnela et al. 2007 Villanueva-Felez, Woolley and Cañibano 2015; J. Wang 2016
Breadth of discussion	(Marsden and Campbell 1984; Mcguire and Bielby 2016
Longevity	Mcguire and Bielby 2016; Petersen 2015
Degree of mutual trust	(Daly and Haahr 2009; Levin and Cross 2004 Marsden and Campbell 1984; Mathews et al. 1998
Degree of friendship	Villanueva-Felez, Woolley and Cañibano 2015
Degree of reciprocity	Granovetter 1973; Hobart, Perlman and Duck 1988 Wellman 1982
Recency	Daly and Haahr 2009; N. Lin, Dayton and Greenwald 1978
Sociability	Mcguire and Bielby 2016; Mitchell 1987
Provide support	**Echebarria2013LimitsRegions**; Mitchell 1987 Wellman 1982; Wellman and Wortley 1990
Prestige difference	N. Lin, Ensel and Vaughn 1981; Marsden and Campbell 1984
Educational difference	N. Lin, Dayton and Greenwald 1978; Marsden and Campbell 2012

Based on the definition of tie strength, Granovetter suggested two kinds of ties – strong ties and weak ties. In recent years, authors have tried to classify the ties into more distinct kind of ties like latent ties (Haythornthwaite 2002), dormant ties (Walter, Levin and Murnighan 2015), intermediate ties (Retzer, Yoong and Hooper 2012). However, these distinctions still fall under the broad spectrum of strong ties and weak ties. Hence, have not been discussed as separate kind of ties in this dissertation.

In general, strong ties are people who you trust and whose social circles tightly overlap with your social circle.(Gilbert and Karahalios 2009; Granovetter 1973) In the professional context, the strong ties might be people with whom you work in a project or in the same group, exchange frequent information about work tasks, and ask for personal advice. (Wu, DiMicco and Millen 2010) In the personal context, it may be the people with whom you have a long relationship history, interact regularly and share every major and minor life experiences.(Granovetter 1973; Krackhardt 1992; Wu, DiMicco and Millen 2010). Family members and close friends are some common examples of strong ties. Strong ties provide emotional support and are more stable and easy to rely upon. In a professional context as well, people rely on their strong ties for protection and comfort in situations of uncertainty. (Granovetter 1973; Krackhardt 1992; Krackhardt and Stern 1988) Organizations also rely on their strong ties during difficult times (Krackhardt 1992; Krackhardt and Stern 1988). Thus, strong ties are useful in a variety of situations.

Weak ties are people with whom you merely have an acquaintance or have had a distant and infrequent interaction with. In professional context, it might be people who work near you but not with you, with whom you may have some casual banter or who are part of the same professional organizations.(Wu, DiMicco and Millen 2010) In the personal context, these might be people whom you may have met at some event like a friend's party.(Gilbert and Karahalios 2009; Granovetter 1973) The weak ties in many cases provide access to novel information, access to information not circulating in your strong ties social circle, help in diffusion of new ideas and also provide new knowledge. (Gilbert and Karahalios 2009; Granovetter 1973; Levin and Cross 2004) Hence, weak ties have been found to be useful in a lot of different situations.

2.1.2 Important theories related to identification of social ties

Prior to delving into the topic of identification of social ties, it is important to understand the meaning of a network. According to Borgatti and Halgin 2011, a network consists of a set of actors or nodes along with a set of ties of a stated type (like friendship) that link them. The choice of a set of nodes and a type of tie chosen by research define a specific network. It is actually dictated by the research question and one's explanatory theory.

A significant aspect of the networks is to bridge the local and global – to offer an explanation that simple processes at the level of individual nodes and links can have complex effects on the network as a whole. (Easley and Kleinberg 2010) Though the two specific theories Strength of Weak Ties and Structural Hole theory are different in the explanation they offer about the reason different ties are useful; at a broader network level, they both rely on the twin notion that the network structure and the position of the node play fundamental roles. Based on these commonalities, it can be seen that both the theories of Strength of Weak Ties and Structural Hole Theory have small differences in ornamentation but are based on how networks work (Borgatti and Halgin 2011). Hence, the two theories help in identification of different ties and can be used in conjunction to identify different kind of ties. In this dissertation, most of the studies have taken a theoretical perspective presented by Granovetter. These two theories are presented in brief in the subsections below.

2.1.2.1 Strength of weak ties theory

During the 1960s, Mark Granovetter conducted a study to understand how individuals looking for a new job found information about these new job vacancies. From his study, he found that most of these individuals found information about the new job through personal contacts. However, these personal contacts were not their close friends but were individuals with whom they had acquaintances.(Easley and Kleinberg 2010; Granovetter 1973) Granovetter proposed a reason for this observation by proposing the strength of weak ties theory.

According to Granovetter 1973, no previous sociological theories had attempted to link micro-level interactions (such as interpersonal relations) with macro-level phenomena (such as social mobility, community organization, and political structure). This strength of weak ties theory was a partial attempt to bridge this micro-macro gap. The theory was limited to small-scale interactions- the strength of interpersonal ties and how the network analysis of this aspect could link to the largely varied macro phenomena like diffusion, social mobility, community organization, and political structure in general.(Granovetter 1973; M. Granovetter 1983). The 'strength' of an interpersonal tie needed to satisfy the following definition: "the strength of a tie is (probably linear) combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie".(Granovetter 1973) This definition is commonly referred to as the definition of tie strength. According to this definition, Granovetter proposed three possible situations between two individuals: the tie was strong; the tie was weak; or the tie was absent. (Granovetter 1973; M. Granovetter 1983) This resulted in two kinds of ties – strong ties and weak ties. Granovetter explained how this idea of tie strength was useful in establishing the micro-macro bridge by giving the following example: Consider two people A and B and a set of people S = C, D, E... with ties to either or both of A and B. The stronger the tie between A and B, the larger the proportion of people in set S that A and B will be tied to that is have a strong or a weak tie. (Granovetter 1973) Thus providing the state of the overall network at a macro level.

In order to explain the findings of Granovetter's study related to finding a new job, he used the concept of strong triadic closure property or the forbidden triad. He based this property on the theory of cognitive balance (Heider 1958) and the general property of triadic closure (Rapoport 1953). According to the property of triadic closure "If two people in a social network have a friend in common, then there is an increased likelihood that they will become friends themselves at some point in the future" (Rapoport 1953). Based on these properties the concept of strong triadic closure was defined as: if A and B have strong ties and also B and C have strong ties then A and C will at least have a weak tie (Easley and Kleinberg 2010; Granovetter 1973). Granovetter used this idea of strong triadic closure property to explain that only weak ties could act as the bridges between two or more different groups. Bridge can be defined as a line in a network which provides the only path between two nodes. The identification of the different kind of social ties could be done by measuring the tie strength. (Easley and Kleinberg 2010; Granovetter 1973; M. Granovetter 1983)

2.1.2.2 Structural hole theory

According to Easley and Kleinberg 2010, a structural hole in an organization is "the 'empty space' in the network between two sets of nodes that do not otherwise interact closely." Structural holes appear in literature extensively in different forms, often in the context of social capital and creativity or the creation of new knowledge (Burt 2004). An individual bridging a structural hole is able to increase his/her social capital through accumulating non-redundant information from varied sources (Burt 1992; Burt 2004).

The concept of structural hole was originally proposed by Ronald Burt to help explain the origin of social capital. Structural hole theory is also based on the property of network closure i.e. most social structures tend to be characterized by dense clusters of strong connections. The structural hole theory also depends on a basic assumption that the homogeneity of information, new ideas, and behavior is in general higher within any group of people as compared to that in between two groups of people.(Burt 1992; Burt 2004) An individual who acts as a bridge between two or more closely connected groups of people can gain significant comparative advantages. This bridging position allows him or her to act as a gatekeeper of valuable information from one group to another.(Burt 1997) Additionally, it also provides an opportunity to combine all the different ideas he or she receives from various sources and combine them to come up with most innovative idea among all (Burt 2004). On the other hand, being a broker between different groups can be difficult at times, as maintaining ties with disparate groups can be fragile and time consuming to maintain (Burt 2001). Thus, the position of the node in the social network determines the kind of social tie that node is. This notion helps in identifying the different kinds of social ties from the network.

2.2 Social ties from social data

2.2.1 Social data and its significance for social ties identification

The advent of the digital era has led to the availability of many different kinds of data. The illustration shown in Figure 2.1 shows one such logical division of the available digital data which is relevant in the context of this dissertation.

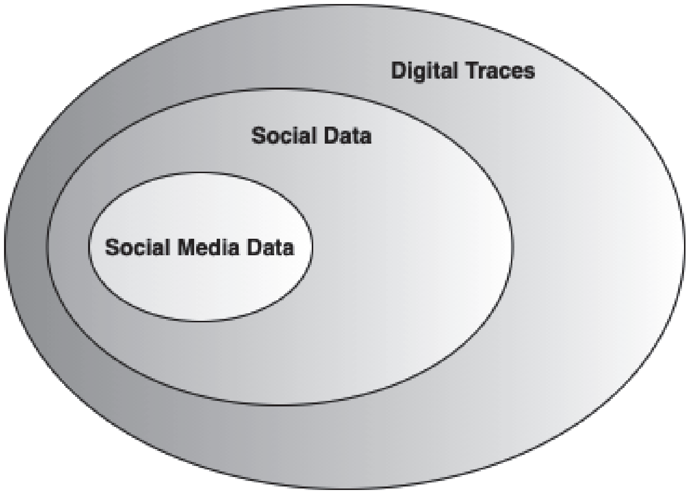

Figure 2.1 Logical division of available digital data

The outermost layer is the data related to digital traces or the digital traces data. The digital traces data can be defined as records of or about activity (trace data) undertaken through an online information system (thus, digital) and can be collected from myriad of technical systems like sensors, phone apps, websites, social media platforms, smart- phone apps, wiki logs, application logs and data from many other digital sources (Howison, Wiggins and Crowston 2011; Stier et al. 2020).

The next layer in the Figure 2.1 is social data. Social data can be defined as the data that typically comes from social software, which acts as an intermediary or a focus for a social relationship (Olteanu et al. 2019; Schuler 1994). It includes data from a variety of platforms—like for social media (e.g., Facebook), question and answering (e.g., Quora), networking (e.g. meetup.com) or collaboration (e.g., Wikipedia)—and for different purposes from finding information (White 2013) to keeping in touch with friends (Lampe, Ellison and Steinfield 2008)). The innermost layer in Figure 2.1 is social media data. Social media data refers to all the data which can be collected from different social media platforms (e.g. Twitter, Facebook, Linkedin) usually using the API of the specific social media platform.

Social data has opened unprecedented opportunities to answer significant questions in different domains about society, policies, and health, has been recognized as one core reason behind progress in many areas of computing (e.g., crisis informatics, digital health, computational social science) (Crawford and Finn 2015; Olteanu et al. 2019; Tufekci 2014). The analysis of social data is believed to provide insights into both individual-level and large human phenomena, with a plethora of applications and substantial impact (Dumais et al. 2014; Harford 2014; Lazer et al. 2009; Tufekci 2014).

Like other areas of research, social data has had a significant impact on the research related to identification of social ties. Since, the social data is inherently related to social relationships, it can be used to develop new approaches, methods and insights about the different aspects of social ties including their identification from different platforms (e.g. Deri et al. 2018; Y. Lu et al. 2016; Petroczi, Bazsó and Nepusz 2007).

In the past decade, the rise in popularity of social media platforms has given rise to new ways to detect, establish, and manage ties online. The availability of social media data over the past decade has enabled researchers to carry out various studies and develop different models for the identification of social ties from social media data (e.g. Backstrom and Kleinberg 2013; Gilbert and Karahalios 2009). Hence, social data and social media data have had a significant impact on the research related to social tie identification. In this dissertation, we have used social media data, more specifically from Facebook and Twitter to carry out the different studies.

2.2.2 Major existing models for identification of social ties from social media

As highlighted in the previous section, the rise in the use of social media platforms has provided opportunities for researchers to utilize social media data for the identification of social ties. In this section, the major existing models which exist for the identification of social ties from social media are highlighted. These models were selected based on satisfying certain criteria.

The following criteria were considered while identifying major models for the identification of social ties from social media. The first selection criterion was that the model for the identification of social ties should be based on the original definition of tie strength and the original concept of strong and weak ties. Hence studies like De Meo et al. 2014 were excluded as this study changed the basic definition of tie strength. The second selection criterion was that the model should have an independent mathematical formulation. Third only models which use social media data for the identification of social ties were to be considered. Thus, for example, tie strength models by Onnela et al. 2007 and by Mattie et al. 2018 were excluded since they used call log data and not social media data. Finally, the model should not just be a minor modification of an older existing model for the identification of social ties from social media data. For example, the model by Fogués et al. 2013 was excluded as it was just a minor modification of Gilbert and Karahalios 2009 model. Finally, only those models for the identification of social ties from social media data were considered which had been cited by other studies. Based on these selection criteria the following existing models for the identification of social ties from social media were selected.

One of the earliest and most commonly cited model for the identification of social ties from social media data was developed by Gilbert and Karahalios 2009. It was one of the earliest model which tried to use social media data to evaluate and predict the tie strength between social media users. Based on the definition of tie strength provided by Granovetter 1973, tie strength has different dimensions that can be operationalized using various measures and predictors. In this model, the different measures and predictors were operationalized by using the various features and functionalities which were available on the social media platforms. Some of the features used for the operationalization of the predictors were specific to just Facebook. This model used a total of 74 measures and predictors operationalized based on the features and functionalities of the social media platforms. The social media data which was used to create the model was collected by crawling and scraping the Facebook page and profile data of the study participants. This model utilized the social media data which included the explicit relationship data of Facebook friends, user profile data, and other explicit Facebook friendship-related data of study participants to calculate the tie strength between the study participants and their online friends on the social media platform. The final outcome of this study was a model which was a linear combination of the predictive variables, the pairwise interaction of the predictive variables, and the network structure where the network structure was based on the Facebook friendship data that was crawled from Facebook. At a later stage, this model was modified to use Twitter data for the identification of social ties. A lot of predictors and measures derived from Facebook data which were used in the previous model were substituted with similar predictors from Twitter data. However, this Twitter data-based model did not have predictors related to all the different dimensions of tie strength (Gilbert 2012).

Around the same time as the model was developed by Gilbert and Karahalios

2009, another model for identifying social ties from social media data based on supervised machine learning was developed by Kahanda and Neville 2009. This model was able to perform the binary task of predicting whether or not a relationship was strong which enabled the identification of strong and weak social ties. This model was built using a publicly available social media dataset - Purdue Facebook network. However, this data is no longer accessible due to the changes in the Facebook terms of service. The Facebook data used in building this model was divided into four different kinds of graphs: friendship graph, wall graph, picture graph, and group graph. These graphs were used to construct a total of 50 different features that were then used for classifying relationship as being strong or not. This was the first model which used supervised learning to identify social ties from social media data.

An unsupervised machine learning model was built by Xiang, Neville and Rogati 2010 for the identification of social ties from social media data. This model used the principle of homophily i.e., people tend to form relationships with similar kind of people. Thus, this model assumed that the stronger the relationship, the higher the similarity. The social media data used in building this model was divided into four kinds of graphs: friendship graph, top-friend graph, wall graph, and picture graph. These graphs were used to infer relationships based on the profile similarity and interaction activity between the social media users of this study. This model was used on two different social media datasets. The first dataset was the publicly available dataset - Purdue Facebook network. However, this data is no longer available due to the change in Facebook terms of service. The second dataset was a proprietary dataset from Linkedin.com.

Another model using logistic regression was developed by Jones et al. 2013 for the identification of social ties from social media data. This model used Facebook data related to measurable online behavior and demographics of the users of the study. Logistic regression provides the probability based on the provided data features, which in the case of this model was to give the probability about whether two Facebook users of the study were close friends in the real world. The developed model could successfully differentiate between the closest friend from the not closest friends in real-world relationships. The model was based on the feature called summed interaction. This feature was the sum of all the interactions between the users. The social media data for this study was collected using the Facebook Id of the study participants and also included some private user data like pokes, and messages. However, this study did not specify if the data was collected using the Facebook API or whether the data was crawled from Facebook.

Over the years, more models for the identification of social ties from social media have also been developed. A lot of these models are also listed in the review paper on tie strength models by Liberatore and Quijano-Sanchez 2017. However, other than the models about identifying social ties from social media data mentioned above, all the other social tie strength models are based on the minor modification of the earlier existing models. For example, the model by Fogues et al. 2018; Fogués et al. 2013 and social tie strength model by Liberatore and Quijano-Sanchez 2017 are based on the minor modification of the original model by Gilbert and Karahalios 2009. Thus, based on the existing literature related to the models for identifying social ties from social media data, the models mentioned above are the current state of the art and can be considered representative of all the existing models for identifying social ties from social ties media data.

2.2.3 Classification of social ties based on the kind of social media data used

In the previous section, the major existing models for the identification of social ties from social media data were presented. Based on the description of these models presented in the earlier section, certain salient features related to the kind of social media data that was used in developing these models can be observed. The models by Gilbert and Karahalios 2009, Kahanda and Neville 2009, and Xiang, Neville and Rogati 2010 used explicit social media relationship data(e.g. Facebook Friends list) in order to develop their models. Also, the models by Jones et al. 2013 and by Gilbert and Karahalios 2009 relied on a lot of private social media data of the users. Both the private user data and the explicit relationship data are no longer accessible from many social media platforms. Thus, there was a need to classify online social ties based on the kind of social media data that was used to identify the social ties.

Social media networks built using social media data can be broadly classified into two kinds - explicit social networks and implicit social networks (Zhou, Duan and Piramuthu 2014) and ties identified from these networks can be classified as explicit social ties and implicit social ties. Explicit social ties are formed on a social media platform when the user explicitly adds other individuals to their network on the social media platform (Reafee, Salim and Khan 2016), like in Facebook, a user adds an individual as a friend or on Twitter, a user follows another account. Hence, explicit social networks are formed using the explicit relationship data available and provided by social media platforms. On the other hand, implicit social ties refer to all the social ties derived from social media data provided by the social media platform excluding the data related to explicit relationships data. Thus, the implicit social networks can be created using a variety of social media data like textual data, conversation data, user profile data, aggregated relationship data, and any other manner of network creation to study the user behavior using social media data (Zhou, Duan and Piramuthu 2014).

In the context of this dissertation, implicit social ties refer to all social ties which can be derived from implicit social networks. This definition is based on the kind of social media data that is used for the identification of social ties. Implicit social ties can be either strong or weak based on the characteristics of social ties defined by Granovetter 1973 in his seminal paper 'Strength of Weak Ties'. In this dissertation, we use the same definition of strong and weak ties as defined by Granovetter 1973.

2.2.4 Significance of bigness of social data towards identification of social ties

Since, the coining of the phrase "big data" many different definitions have been provided for it (e.g. Harford 2014). Big data science can refer to a field that processes and manages high-volume, high-velocity and high-variety data in order to extract reliable and valuable insights (Olshannikova et al. 2017). This buzz around big data has also led to a large interest to include the 'social' aspect of the big data. The two most common terms used in this context are big social data and social big data. Big social data has been defined as any high-volume, high-velocity, high-variety, and/or highly semantic data that is generated from technology-mediated social interactions and actions in a digital realm, and which can be collected and analyzed to model social interactions and behavior. (Olshannikova et al. 2017) On the other hand, Ishikawa 2015 defines social big data as "analyzing both physical real world data (heterogeneous data with implicit semantics such as science data, event data, and transportation data) and social data (social media data with explicit semantics) by relating them to each other, is called Social Big Data science or Social Big Data for short".

Both the definitions - big social data and social big data have the 3V's in common i.e. volume, velocity and variety. Here volume refers to the exponential growth of social data while velocity is related to the fact that social data is generated and distributed with tremendous speed. Variety is related to various types and forms of social data sources: structured, semi-structured or unstructured. Variety can also mean the difference of formats (for example, text, image, video).(Olshannikova et al. 2017) However, Ishikawa's definition of social big data also includes a fourth V which is related to vagueness. According to Ishikawa, bigdata inherently contains vagueness due to combination of various types of data to be analyzed, which lead to inconsistency and deficiency. It is also related to the issues of privacy and data management as social data involves individuals' personal information. (Ishikawa 2015)

In the previous section, it was highlighted that the increase in the data access restriction of social media platforms has made it impossible to use the existing models of social ties identification based on social media data. In the past, it was possible to directly access the explicit relationship data which made it easy to even use small amounts of social media data for the identification of social ties based on social media data. However, in order to make implicit networks, logically there is an inherent need to have at least a larger variety of social data and not necessarily a larger volume of data. This is due to the fact that different kinds of implicit social networks will be needed to identify the social ties more accurately from the social data. Having access to and using a large variety of social data for developing new approaches and methods for the identification of social ties from social data will enable the reduction of vagueness associated with big data. Hence, in this dissertation, a conscious effort was made to use a large variety of social data and not just a large volume of social data in order to develop approaches and methods which can identify social ties from social data and big social data.

2.3 Challenges in identification of social ties from social data

Increase in data access restrictions and the need for implicit social ties

During the initial years of the existence of social media platforms, there were not many standardized ways and limitations on who and how social media data could be collected. It meant that methods like data scrapping were perfectly legal ways of collecting social media data (e.g. Gilbert and Karahalios 2009). However, with the rapid growth of social media platforms, it became important to maintain the data integrity and data access of social media platforms. In order to achieve this initial objective, different social media platforms adopted the use of authenticated API (application programming interface). These APIs achieved the initial goal of data integrity and data access management.(Hogan 2018) However, the growth and proliferation of these social media platforms resulted in these social media platforms turning into one of key players in the growing data markets. The business models of these companies moved towards providing privileged access to social media data, and the resulting valuable insights could be gained from this user-generated social media data.(Hogan 2018; Weller et al. 2013) Beyond this change in the business model of social media platforms, the introduction of new data regulations (such as GDPR) and the overall push towards more privacy has further reduced the data access from social media platforms. The social media platforms in general are moving towards more restrictive data access policies.(**guscientific**)

This shift towards restrictive data access policies of the different social media platforms, can severely impact how social ties can be identified from social media data(Gupta, Kärkkäinen, Torro et al. 2019). Some of these challenges related to social data are listed below:

1. There are severe restrictions on the volume and time period for which the social media data can be accessed from the social media platforms.
2. It is almost impossible to get access to explicit relationship data and user's private data.
3. Either the access to historical data is severely restricted, or there is no access.
4. Current models for identifying social ties based on social media data use explicit relationship data and user's private data.
5. Data can be accessed from social media platforms using only the data API of the platforms. Another means of collecting data (scraping the data) is not allowed and is illegal in most cases.

This increase in the restrictive data policies of social media platforms has resulted in the need to assess the utility of the existing social tie identification models based on social media data. The analysis of the existing social tie identification models based on social media data shows that these models relied on using a lot of explicit relationship data and also methods like data crawling which are no longer allowed and the explicit relationship data is impossible to get in most cases (Gupta, Kärkkäinen, Torro et al. 2019). These changes in the availability and the kind of the social media data has led to the need to explore other possible ways of identification of social ties. One of the important means of overcoming these challenges is to rely on publicly available social media data and use the notion of implicit social ties to identify the different kinds of social ties. Since, the implicit social networks do not use any explicit relationship data, it is possible to create implicit social networks based on the available social media data and use these networks for identification of different kinds of implicit social ties.

2.3.1 Challenges related to social data and big data for identification and use of social ties

As highlighted in the previous section, there are particular challenges in using social data to identify and use social ties. Many of these challenges are related to the restrictive data policies of the different social media platforms. At the same time, when using large social data to identify social ties, specific challenges inherent in working with big data are also present. Big data has been defined using many different versions of 'V's. One of the popular definitions by Gandomi and Haider 2015 of big data, which takes into 6 'V's of big data, is used to understand the different challenges encountered when using big data to identify social ties. These challenges are presented below:

- The first 'V' is related to the definition of big data is related to Volume. This refers to the magnitude of data. Identifying social ties can be challenging as the large volume of the social data may make it challenging to filter out the relevant social ties from the data. This situation is particularly common in the case of using publicly available social data to identify implicit social ties.
- The second 'V' is related to the high level of a variety of data. The data can be from different sources like sensor data, social media platforms, clickstream data, bibliographic data, and various other sources. This also includes the

structural heterogeneity in a dataset. The data can be structured or unstructured. Most social data involve Text, images, audio, and video, which are examples of unstructured. Unstructured data sometimes lacks the structural organization required by machines for analysis. Thus, even in identifying social ties from social data, there is a challenge to use and develop new approaches and methods that make it possible to carry out the data analysis.

- The third 'V' is related to velocity. It refers to the rate at which data is generated and the speed at which it should be analyzed and acted upon. This aspect of big data may be critical for the successful development of real-world applications like conference recommendation systems that would incorporate the identification of social ties while giving out recommendations.
- The fourth 'V' is related to the veracity of the data. This represents the unreliability inherent in some sources of data. This is especially true in the case of social data. For example, the sentiment analysis results of social media data are uncertain in nature as they entail human judgment. However, at the same time, the results of the analysis contain valuable information. Thus, while using social data to identify social ties, it is essential to consider the veracity aspect of big data. Especially in the case of identifying implicit social ties from publicly available data, the aspect of veracity is critical as, unlike explicit relationship data used in earlier models, implicit ties do not use any explicit relationship data.
- The fifth 'V' is related to variability, which is related to the variation in the data flow rates. This also includes the aspect of complexity as big data is generated from a myriad of data sources. This creates a critical challenge: the need to connect, match, cleanse and transform data received from different sources. This aspect of big data may also be critical for the successful development of real-world applications like conference recommendation systems which would incorporate the identification of social ties while giving out recommendations.
- The final 'V' is related to value. In big data, the data received in the original form usually has a low value relative to its volume i.e., it has 'low-value density'. Thus, in the context of identification of social ties, it could result in the need to have a large volume of data even to identify a small number of social ties. This is very common in the case of identification of implicit social ties

as there is no direct access to explicit relationship data which can be used for identifying the social ties.

This subsection has presented some of the challenges related to the use of social data and big data to identify social ties. In this dissertation, the different studies were designed and conducted to address some of the challenges presented above.

2.4 Various contexts for making use of social ties

The role of social ties both strong ties and weak ties was originally used by Granovetter to study the role of the different kind of social ties in the context of job seeking. Over the decades, the role of social ties has been analysed in many different fields ranging from social science to computer science, economics and many other areas (Gilbert 2012). The role of social ties has been analyzed in multiple different context like:

- innovation and creativity (Sosa 2011)
- knowledge transfer Levin2004
- information diffusion (Gilbert and Karahalios 2009)
- content sharing (Zhan Shi, Huaxia Rui and Whinston 2014)
- social influence (Aral and Walker 2014)
- finding romantic relationships (Backstrom and Kleinberg 2013)
- group movie recommendations (Liberatore and Quijano-Sanchez 2017)
- cyberaggresion (Wegge et al. 2015)
- socially enhanced applications (Servia-Rodríguez et al. 2014)
- privacy assistance (Fogués et al. 2013)

At the same time the role of social ties has been extended from its original use at an interpersonal level to organizational and inter-organizational levels as well (e.g. Walter, Levin and Murnighan 2015, Wiese et al. 2015). Hence, it can be seen that there are multiple fields and contexts in which the role of social ties has been studied and can be further analyzed. However, in order to limit the scope of this dissertation and gain a detailed understanding, in this dissertation the role of social ties has been studied in two different contexts - events, and crowdfunding. Before finalizing these two contexts for carrying out the research studies, there were attempts to use data from other contexts and experiment with them but they did not materialize, due to many different constraints (like access to relevant data). Several different factors led to the selection of two specific contexts - events and crowdfunding. These different factors are explained in this section.

An essential practical factor that favored the selection of these contexts was the involvement in the ongoing research projects, which were the projects - COBWEB and BigMatch.The project aimed to enhance knowledge work and co-creation using Big Social Data Analytics. One of the research contexts for these projects was the context of events. This provided the impetus and easier access to the data needed to carry out some studies related to identifying implicit social ties from publicly available social media data in case of events.

Another crucial practical factor that enabled researching these contexts was the research partners involved. These research partners had specific competencies that enabled access to the data and the knowledge of the different kinds of analysis that could be carried out. For example, the Center for Business Data Analytics (cbsBDA) at the Department of Digitalization, Copenhagen Business School, was a research partner involved in these projects. They provided access to the Social Data Analytics Tool (SODATO) (Hussain and Vatrapu 2014) which enabled access to data from open Facebook groups or open Facebook walls. This tool enabled getting access to large datasets related to open Facebook groups and walls. Hence, the partners involved in the research also impacted the selection of the research contexts.

The social data is critical for carrying out the research studies related to this dissertation. Thus, it was essential to select contexts where the role of social data was vital for studying the underlying phenomena. Based on the earlier studies, it was established that social media played a significant role in the contexts of networking in events and crowdfunding success. Thus, there were complimentary synergies related to the role social ties could play in these two different contexts. This helped in selecting these two research contexts for this dissertation.

At the same time, the reasons for the identification and use of social ties were very different in both these contexts. In the case of events, the social ties were predominantly used to get access to new social connections, which could be helpful in the future. In contrast, in crowdfunding, social ties were used to spread the reach of a crowdfunding project campaign and enable receiving more funding. These fundamental differences in the role social ties play in these two contexts made them interesting to further research.

In the subsequent subsections, details related to these two research contexts and their importance, and relevance for research related to social ties in general and more specifically to this dissertation are provided. These subsections explain the novelty, relevance, and need for carrying out further studies.

2.4.1 Networking in events

In this subsection, the importance of the context of events and the important role social ties play in the business events (like conferences) is discussed. Events (like business conferences) play an important role in transferring scientific, managerial, and other types of information and knowledge among the conference attendees or participants.

2.4.1.1 Motivation of event participants for attending an event

There are multiple expectationsthath an event participant has for attending an event. According to Severt et al. 2007, there are different motivations of the conference participants in attending a conference or an event. These motivations are education benefits, products and deals, activities and opportunities, networking, and the convenience of the conference.

The motivation related to the educational benefits refers to the diverse possibilities a participant may have access to by attending the conference. These opportunities include interesting conference programs, different educational information, career enhancement opportunities, and educational information at exhibits.(Severt et al. 2007) The educational benefits have also been found to be one of the top motivators for attending a conference.(Ross et al. 2011; Severt et al. 2007)

The motivation related to the product and deals factor may not apply to all kinds of conferences and is more relevant in the case of conventions. This factor refers mainly to the products available for purchase at exhibits and the deals on conference packages.(Severt et al. 2007)

The motivation related to activities and opportunities is associated with business opportunities, association-related activities, touring opportunities, visiting friends or relatives, or attending a guest program.(Oppermann and Chon 1997; Severt et al. 2007) The factors like traveling opportunities, visiting friends, or relations have been found to have minimal influence on the motivation for attending a conference.(Severt et al. 2007)

The motivation related to networking is about opportunities to meet new people who may share similar interests or provide relevant information. The participants use these conferences to establish connections with new participants who may be potentially helpful in the future.(Oppermann and Chon 1997; Severt et al. 2007; A. Zhang, Bhardwaj and Karger 2016) Networking has been one of the most significant motivations for attending conferences in many different studies. (see examplеOppermann and Chon 1997; Ross et al. 2011; Severt et al. 2007

The motivation factor related to convenience of the conference includes aspects like reasonable travel time to the conference location, reasonable conference pricing, and the chance to integrate the conference into the participant's work schedule easily. These factors sometimes play a role in the overall motivation of a participant to attend a conference. (Severt et al. 2007) For example, many potential conference participants may opt-out of an educational conference organized during summer vacations due to the inconvenience in the scheduling.

From the above-mentioned different motivators for attending an event (like a conference), it has been found that networking and the opportunity to gain educational benefits are the most important factors. (Oppermann and Chon 1997; Ross et al. 2011; Severt et al. 2007; A. Zhang, Bhardwaj and Karger 2016) Thus, it can be seen that the opportunity to network is one of the most important motivations for attending a business event.

2.4.1.2 Common ways of networking in an event

Currently, there are three different methods of identifying and networking with potentially useful contact at events (like conferences). The first way is related to a chance encounter with a relevant person. The second method is based on the structure or how the conference organizers organize the conference. The final method is based on the use of technology-based interventions like conference recommendation systems.

Many times, event participants end up meeting other relevant new participants by chance. In such meetings, there are no structured interventions based on the organizers' efforts in organizing the conference or based on the conference recommendation systems. However, such serendipitous encounters are minimal and cannot be planned by the conference participants in advance.(Reinhardt, Maicher et al. 2011; Ross et al. 2011; A. Zhang, Bhardwaj and Karger 2016). Thus, this method of chance encounter is not very helpful in increasing the chances of a conference participant meeting another potentially helpful contact.

In many events (like conferences), event organizers use past participant feedback and the analysis of the previous events data to design the layout for subsequent conferences, promoting networking among conference participants. Other event organizers related interventions include organizing common transportation to the event venue and physical spaces; and seating order based on the common themes the event participants are interested in or relevant to their learning objectives. For example, the seating arrangements around the lunch tables and coffee tables can be based on the identified topics of interest.(Aramo-Immonen, Jussila and Huhtamäki 2015; Aramo-Immonen, Kärkkäinen et al. 2016) Thus, in many cases, the event organizers' different layout arrangements improve the possibility of the event participants finding valuable contacts.

The need to meet relevant people at an event has given rise to the development of technology-based interventions like conference recommendation systems (eg. brella.io, confer). These conference recommendation systems can provide suggestions to meet certain other participants that may be potentially useful.(Hornick and Tamayo 2012) Most of these systems have relied on giving recommendations based on certain keywords usually extracted from the even participants' registration form or some other information that the participant had provided at the time of event registration (Hornick and Tamayo 2012; Zhong, Yang and Nugroho 2015). In recent years, some studies have tried to incorporate other sources of data like bibliographic data, co-occurrence data, participant's mobile device data, and also data from sites like epinions.com, Flickr to provide more relevant recommendations(Gupte and Eliassi-Rad 2012; A. Zhang, Bhardwaj and Karger 2016; Zhong, Yang and Nugroho 2015). Hence, technology-based systems are also being used to provide ways to meet potentially valueable contacts at a conference.

2.4.1.3 Role of social media in events

The arrival of social media has allowed the traditionally passive role of event participants and delegates to be more active. Social media has provided novel ways for integrating participants in pre-event planning, during-event participation, and after-event activities, such as sharing event-related information. This has also allowed the creation of tailored event content and more helpful networking opportunities.(Jussila et al. 2013; Reinhardt, Ebner and Beham 2009)

Social media provides an informal way to interact with other conference participants and acts as a secondary communication route. Some of the earlier studies have found that social media is used to maintain existing relationships and establish new relationships (Ahn and Park 2015; Boyd, Golder and Lotan 2010). This is true in the case of conferences as well. The need to establish new social relationships and maintain existing relationships were among the top reasons for using social media in an event setting (Reinhardt, Ebner and Beham 2009; Ross et al. 2011).

Many studies have also explored the use of social media by conference organizers. Event organizers have used social media to crowdsource certain event activities like event marketing and other co-creational activities to organizers and participants alike (e.g., Jussila et al. 2013; Ross et al. 2011). The organizers and presenters have used social media to facilitate instant discussions during presentations. (Aramo-Immonen, Jussila and Huhtamäki 2015; Jussila et al. 2013) The visualization of the social media discussion data related to the conference has been used to identify the most influential participants and identify the conference's common important themes (Aramo-Immonen, Jussila and Huhtamäki 2015). These visualizations can help the organizers design better layouts, which can be helpful in effective networking for the event participants (Aramo-Immonen, Jussila and Huhtamäki 2015; Aramo-Immonen, Kärkkäinen et al. 2016).

Thus, it has been found from prior studies that both the event participants and event organizers use social media for multiple purposes. One of the most important reasons for using social media is for networking.

2.4.1.4 Relevance of social ties for enhancing networking in events

From the earlier studies, it is known that one of the primary aims of event participants in an event is to meet new potentially beneficial other participants. Nevertheless, there is minimal time in such events for networking opportunities.(Oppermann and Chon 1997; Severt et al. 2007; A. Zhang, Bhardwaj and Karger 2016) This shortage of time serves as a potential challenge and restricts the chance of meeting potentially valuable contacts. In order to avoid this challenge, some studies have shown that participants use social media as a channel for networking (Ross et al. 2011). As highlighted in the above section, the use of social media for networking and information seeking is one of the most important reasons for social media use in an event setting (Ebner et al. 2010; Reinhardt, Ebner and Beham 2009; Ross et al. 2011). Thus, intuitively, it would make sense to use social media data to identify social ties even in the event setting.

However, the existing models that use social media data to identify social ties have relied on using explicit relationship data and private user data from social media platforms. (Gupta, Kärkkäinen, Torro et al. 2019) This can be a potential challenge in case of events. Such an approach would require access to the event participants' personal social media data, which is almost impossible to access in most cases. Hence, the already established methods for identifying social ties from social media data may not be directly helpful in an event setting. On the other hand, it may be easier to use publicly available social media data about an event to identify social ties.

The current technology-based interventions like conference recommendation systems that enable networking in events use different strategies like matching keywords from event registration data (e.g., Hornick and Tamayo 2012; Zhong, Yang and Nugroho 2015). However, these systems do not yet incorporate or consider the perspective of social ties identified using social media data. The development and incorporation of such an approach may result in better and more relevant recommendations and provide a more efficient way to network and meet useful people in an event like a conference. In order to reach this stage, there is a need to first find ways in which social ties can be identified from social media data in the context of an event like a conference.

2.4.2 Crowdfunding and social ties

In this subsection, the context of crowdfunding is discussed. The different types of crowdfunding platforms which exist are provided. Also, the role of social ties and social media in the success of crowdfunding projects is discussed.

2.4.2.1 What is crowdfunding ?

Over the past few years, crowdfunding has attracted much attention and gained traction from the public in general as well as from academia. Crowdfunding provides an alternative channel to traditional financing institutions. In the past decade, crowdfunding has become an indispensable and novel means of raising capital to carry out projects which have earlier not been possible easily. According to Mollick 2014, "crowdfunding is a novel method for funding a variety of new ventures, allowing individual founders of for-profit, cultural, or social projects to request funding from many individuals, often in return for future products or equity." Thus, crowdfunding is a type of crowdsourcing mediated by digital platforms that enable entrepreneurs of all types - social, cultural, artistic, or for-profit - to raise capital from the crowd to launch new ventures or causes. (Hong, Hu and Burtch 2018; Mollick 2014) Crowdfunding provides an alternative means of connecting people that need resources to carry out a project or an idea with those willing to anchor them to begin a project or a business through digital platforms (Madrazo-Lemarroy, Barajas-Portas and Labastida Tovar 2019).

2.4.2.2 Types of crowdfunding platforms

The rise in the popularity of crowdfunding has also resulted in the formation of many different crowdfunding platforms that have followed different kinds of business models. These different crowdfunding platforms can be broadly classified into four different categories (Kromidha and Robson 2016). A brief description of each of these categories is presented below:

1. Donation-based crowdfunding platforms: This kind of model is usually used to support charity projects. It helps connect possible donors with the charitable causes of their interest (Y. Zhang et al. 2020). Gofundme.com, milaap.org, justgiving.com are some of examples of donation based crowdfunding platforms.
2. Investment or equity based crowdfunding platforms: In this kind of model, the funders have a chance to invest or buy equity in a business or a project. This kind of crowdfunding model differs from the reward-based or donation-based crowdfunding platforms as the incentive for investment is not related to

receiving a product or directly supporting a cause (Hervé et al. 2019). Some common examples of this kind of crowdfunding platform are Fundable.com, crowdcube.com, and seedrs.com.

3. Crowdsourcing and group lending-based crowdfunding platforms: In this kind of model, the funders expect a return on their donation with or without interest.Kiva.org, and prosper.com are some example of this kind of crowdfunding platform.

4. Reward-based crowdfunding platforms: This is one of the most popular kinds of crowdfunding business models. In this model, funders can get tangible or intangible rewards like a personalized model of the product being developed by a project or a simple 'Thank you' note, for example, after the project finalization. The reward can be a product, artwork, or any reward based on the size of his/her donation (Giudici et al. 2012). Some of the most popular reward based crowdfunding platforms are Indegogo.com and Kickstarter.com.

The reward based crowdfunding platforms can be further categorised into two categories. These two categories are "All-Or-Nothing" (AON) and "Keep-It-All" (KIA). In the "All-Or-Nothing" kind of reward based crowdfunding platforms, the creator of the crowdfunding project only receives the funding amount if the funding goal of the crowdfunding project is achieved. On the other hand, in case of "Keep-It-All" kind of kind of reward based crowdfunding platforms, the creator of the crowdfunding project receives the entire funding amount regardless of achieving the funding goal of the crowdfunding project. Indegogo is an example of "Keep-It-All" kind of kind of reward based crowdfunding platform while Kickstarter is an example of "All-Or-Nothing" kind of reward based crowdfunding platform. (Cumming, Leboeuf and Schwienbacher 2020; Davies and Giovannetti 2018)

2.4.2.3 Different factors related to crowdfunding success

A growing body of literature has investigated the factors that influence the success of crowdfunding projects. The crowdfunding process has been studied by some scholars, but some scholars have looked at the impact of social capital on fundraising (J. S. Hui, Gerber and Greenberg 2012).

Another research stream (Gerber and J. Hui 2013) has focused on investigating the motives and deterrents for participation in crowdfunding platforms. The third stream of literature (Kang, Jiang and C. H. Tan 2017) has focused on investigating the effect of fundraisers' social capital. Furthermore, research in (Mollick 2014) has mainly examine design strategies for crowdfunding projects, including design strategy dealing with designing the project content, reward schema (N. Zhang, Datta and Kannan 2014), and social media usage related to the crowdfunding project.

All crowdfunding transactions include three kinds of actors: the person who creates the crowdfunding project (also identified as the project owner), the person who funds the crowdfunding project (also identified as a project backer), and the crowdfunding platform itself (Belleflamme, Lambert and Schwienbacher 2014; Madrazo-Lemarroy, Barajas-Portas and Labastida Tovar 2019; Mollick 2014; Schwienbacher 2018). Based on these three kinds of actors of the related relevant literature can be categorized. A stream of literature has focused on the relationship between crowdfunding project owner and their role in the success of crowdfunding projects. Several positively affected these studies have concentrated on the extent of the project owners' social networks, the presence of a product video, and project owners' geographical proximity raising the probability of achieving a project's funding goal(Frydrych et al. 2014; Mollick 2014; Saxton and L. Wang 2014). It has also been found that project owners who are successful have been found to have a high number of online friends, relevant backgrounds, and external endorsements (Shneor and Vik 2020). In many cases, project owners' friends and family have positively influenced crowdfunding success (Agrawal, Catalini and Goldfarb 2011; Moritz and Block 2016). Some studies have also found that emotional and cultural factors related to project owners seem to positively influence crowdfunding success (Burtch, Ghose and Wattal 2014; M. Lin and Viswanathan 2016).

Another literature stream has focused on the role of project backers' on crowdfunding success. Research has examined the different reasons behind project backers' funding decisions. In some of these studies, social reputations and intrinsic motivations influenced project funding (Y. Lin and Boh 2020). Several studies have examined the social network effect on project backers' behavior (such as herding) (De et al. 2015; Saxton and L. Wang 2014). Studies have also looked at the different signals and signaling mechanisms that influence the decisions of project backers to fund projects (Gleasure and Feller 2018; Kromidha and Robson 2016; M. Lin and Viswanathan 2016). A recent study investigated how influential project backers and their networks affected crowdfunding success (Y. H. Tan and Reddy 2021). The role of embeddedness, centrality and social influence on project backers' behavior in a crowdfunding network was also examined in a recent study (Chung, Y. Li and Jia 2021).

Yet, another stream of literature has focused on crowdfunding platforms or intermediaries. In some studies, crowdfunding platforms have also been assessed for their ability to reduce information asymmetries (N. Wang et al. 2021) and to compare the regulatory environments in which they operate (Hornuf and Schwienbacher 2017; Klöhn 2018). Another study examined the trust-building mechanisms of crowdfunding platforms (Greiner and H. Wang 2010), while another investigated the role of network effect on crowdfunding platforms' growth (Thies, Wessel and Benlian 2018). Some studies have classified the different kinds of crowdfunding platform archetypes (Kromidha and Robson 2016; Paschen 2017). Some other studies have focused on different kinds of funding models of crowdfunding platforms (Cumming, Leboeuf and Schwienbacher 2020).

From the existing literature, it can be seen that there is no existing research related to the role of implicit social ties towards crowdfunding project success. Also, it can be observed that the role of project owner's social ties in crowdfunding success has been studied to a great extent. However, research focused on the role of project backer's social ties on crowdfunding success is not present.

2.4.2.4 Role of social ties in crowdfunding success

Boosting the number of backers through any means possible has been a key goal since crowdfunding was launched. Project owners have tried to build online crowdfunding communities surrounding their projects using social media, crowdfunding platforms, and online and offline social ties (Gerber and J. Hui 2013). A few studies have investigated the role of project owners' social ties in terms of how their presence or the number of connections on social media can influence the probability of crowdfunding success (Colombo, Franzoni and Rossi-Lamastra 2015; Mollick 2014). Some studies have also shown that the background and closeness of project owners' social ties can significantly impact project funding timing and achieving the funding goal (Agrawal, Catalini and Goldfarb 2015; English 2014) . A few studies have also analyzed the role of project owners' social ties as early project backers and their impact on crowdfunding success (Ordanini et al. 2011; Shneor and Vik 2020) . Other studies have found that project owners' strong ties (family and friends) are crucial for raising crowdfunding funds in various domains such as the arts, culture, academics, and peer-to-peer (P2P) lending. (De et al. 2015; English 2014). Some studies have also used the concept of tie strength to analyze how the strength of project owners' social ties affect crowdfunding success (Borst, Moser and Ferguson 2018; Kim and Z. Zhang 2018). Additionally, research has shown that social ties' geographical location and distance can affect crowdfunding success in varying degrees.(Agrawal, Catalini and Goldfarb 2011; Burtch, Ghose and Wattal 2014; Kang, Jiang and C. H. Tan 2017; Zheng et al. 2014). Thus, different social ties' roles are very important in the context of crowdfunding success.

As highlighted earlier, previous research has revealed that social ties generally influence crowdfunding projects' success (Borst, Moser and Ferguson 2018; Kim and Z. Zhang 2018; Madrazo-Lemarroy, Barajas-Portas and Labastida Tovar 2019). Crowdfunding projects' success has been shown to be influenced by both strong and weak ties. Strong ties generate links between actors in a network, and weak ties act as bridges between members. In crowdfunding, digital platforms (including social media platforms and crowdfunding platforms) can be utilized to strengthen both types of relationships. While strengthening existing social ties, they also help develop new social ties as a campaign evolves, utilizing digital platforms to reach funding objectives (Granovetter 1973; Madrazo-Lemarroy, Barajas-Portas and Labastida Tovar 2019). Therefore, social ties can provide a valuable perspective to understand the different factors affecting crowdfunding success. More specifically, the role of implicit social ties towards crowdfunding success can be analyzed to understand better the factors that affect crowdfunding project success.

As project owners and project backers are the main actors in crowdfunding projects, understanding the role of both parties' social ties on social media is crucial to understanding how social ties affect a project's success in a crowdfunding project. Many studies have investigated the effect of project owners' social ties (friends and family) over the past few years. This research shows that crowdfunding project success and failure are both affected by the project owner's social network.(Madrazo-Lemarroy, Barajas-Portas and Labastida Tovar 2019) Early on in a crowdfunding campaign, friends and family tend to make contributions. However, people who are less familiar with the project contribute later, in reaction to the contributions of the first actors (Agrawal, Catalini and Goldfarb 2015; Colombo, Franzoni and Rossi-Lamastra 2015; Zheng et al. 2014). This dependence may result from project owners possessing additional information about a project and effective social ties that help them make positive choices about their contributions (Agrawal, Catalini and Goldfarb 2015). There has, therefore, been a great deal of literature devoted to the role social ties of project owners play in crowdfunding success. In contrast, despite project backers being among the key players in any crowdfunding project, very little research has examined how their various characteristics influence crowdfunding success. There have been a few studies investigating the effects of crowd participation or funding goal on a crowdfunding project's degree of success (Shneor and Vik 2020; C. Yin, L. Liu and Mirkovski 2019). Recently, a study attempted to analyze the influence of influential backers on crowdfunding projects (Y. H. Tan and Reddy 2021). However, no existing studies have examined the role of project owners' and project backers' social ties (implicit ties) in determining a crowdfunding project's degree of success. Thus, there is an opportunity to analyze the role of implicit social ties of project owners and possible project backers towards crowdfunding project success.

3 RESEARCH METHODOLOGY

The primary objective of this dissertation is to identify implicit social ties from social data and study their role in different contexts. Drawing from the objective of the dissertation, we explain in this section, the decisions made to conduct this research and the methodological and method-related options selected to answer the research questions suitably.

3.1 Research approach

This section clarifies the research approaches taken in this dissertation. It also describes the factors that have most affected the choices of the research approaches and methods. The main data sources used in the different sub-studies are also described. The researcher's choice in practice are affected by the kind of research objective and also access to the data.

Research is generally viewed as a process, where "a set of activities unfold overtime" (Ghauri, Grønhaug and Strange 2020). The research approach is defined as "the plans and procedures for the research that span the steps from broad assumption to detailed method of data collection, analysis, and interpretation" (J. W. Creswell and J. D. Creswell 2017). One of the effective research frameworks which illustrates these different steps is the 'Research Onion' developed by Saunders and Thornhill 2019.

In order to devise a research strategy, especially from the methodological and methods perspective, we have used the research onion framework. The idea of identifying implicit social ties from publicly available social media data was a very new concept when this dissertation began and is still very novel for studying the role of implicit social ties in different contexts. This dissertation intends to initially go into the depth of understanding how implicit social ties can be identified from social data and big social data, and at a later stage understand the role implicit social ties play in some different contexts like crowdfunding. The outer layer of the research onion model is related to the philosophical stance. This dissertation adopted a stance that was somewhere in between positivism (only observable phenomena can provide credible data, and facts) and pragmatism (either or both observable phenomena and subjective meanings can provide acceptable knowledge depending on the research question).

The next layer of the research onion model refers to the approach adopted for theory development. There are three main approaches deductive, inductive and abductive. The deductive approach concentrates on using the literature to identify theories and ideas that the researcher will test using data. The inductive approach involves collecting data and developing a theory based on the results of data analysis. On the other hand, instead of moving from data to theory (inductive approach) or from theory to data (deductive approach), an abductive approach moves back and forth, in effect combining deduction and induction.(Saunders and Thornhill 2019) Based on these definitions, during the initial stages of this dissertation, the inductive approach was adopted where the analysis of the data revealed some interesting propositions for example how the use of a specific social media channel needs to be established before using the social media data for identifying the implicit social ties. At the same time, this research also adopted the deductive approach for example hypothesis developed in one of the sub-study is based on theory and was tested using data. Thus, at a broader level, it can be said that this research used an abductive approach as the research moved back and forth from theory to data or from data to theory.

The subsequent layers of the research onion model refer to the methodological choices and strategies adopted by the research project. In terms of methodological choices, this research primarily used quantitative methods for research design. This dissertation uses multiple research strategies which are case study, survey and experiment.

The initial phase of this research aimed to gain an in-depth understanding of how implicit social ties could be identified using social data and what kind of role these social ties play in the specific context of business. In order to achieve this objective in our research a single-case case study was used. The aim of a case study is to investigate the case thoroughly by using multiple methods and data sources which all aptly reflect the research problem (Ghauri, Grønhaug and Strange 2020; R. K.

Yin 2018). According to R. K. Yin 2018, there are approximately four categories of case study design: a single-case and a multiple-case design with both having either holistic or embedded units of analysis. Single-case design is relevant when the case represents a critical, extreme, or unique case worth documenting; a typical or a revelatory case with unique opportunity to observe a previously inaccessible or common situation; or a longitudinal case, where the same case is studied at various points in time.(Saunders and Thornhill 2019; R. K. Yin 2018) In this dissertation, two different single-case case studies were carried out. The first case study was related to analyzing the social media data of the community of indie game developers. This study will henceforth be referred to as Case: Indie Game Developer Community in the dissertation. The second case study was related to the social media data related to the participants and organizers of an event called CMAD. This second case study will henceforth be referred to as Case: CMAD in the dissertation. A detailed description of both case studies is provided in the later part of this section.

The initial phase of this research was exploratory in nature. The survey strategy is generally associated with a deductive research approach and is primarily associated with exploratory and descriptive research. Designing questionnaire asking 'what', 'who', 'where', 'how much' and 'how many' kind of questions is part of using survey strategy. (Saunders and Thornhill 2019) This kind of questionnaire was needed during the initial phase of this research as it helped in understanding the relationship of social media data based implicit ties against the theoretical framing. However, data collected using survey strategy is usually not as wide ranging as some other data collection strategies (Saunders and Thornhill 2019). For example, a questionnaire can only contain a very limited number of questions and generally it is dependent on the goodwill of the respondent which cannot be presumed. This limitation of survey strategy was also encountered in our research. The survey strategy was used in the Case: CMAD of this dissertation.

In the later phase of the research, the objective was to understand the role of implicit social ties toward some specific business objective in a specific context. This goal was achieved using the research strategy of experiment. Experiment is a kind of research which owes its origin to natural sciences but is also frequently used in psychological and social science research. The purpose of an experiment is to study the change an independent variable causes in a dependent variable. One of the common ways of studying this relationship between dependent and independent variables is to use hypothesis testing. Experiments are commonly used in exploratory and explanatory research to answer 'what', 'how' and 'why' questions. (Saunders and Thornhill 2019) In this dissertation the experiment research strategy was used to understand how implicit social ties have an impact on crowdfunding project success. In order to carry out the related experiments, the crowdfunding project data was collected from one of the largest crowdfunding platforms Kickstarter.com. Along with the crowdfunding project data, the social media related to the crowdfunding projects was collected from Twitter. These datasets were then used to test two different sets of hypotheses related to the role of implicit social ties in crowdfunding success. This experiment is henceforth referred to as Experiment: Kickstarter in the dissertation. The details related to how the different sets of hypotheses related to Experiment: Kickstarter are explained in a later section.

3.2 Case studies and experiment

In this section, the description related to the two different case studies and the empirical experiment which were carried out in order to address the research questions of this dissertation is presented. First, the description related to Case: Indie Games Developer Community is presented. This case study was used for the development of an approach related to the identification of social ties from a large social data dataset. Second, the description related to Case: CMAD is presented. This study was carried out for the identification and role of different kinds of social ties using publicly available data about an event. Finally, the description related to Experiment: Kickstarter is presented. This was done for understanding the identification and role of social ties in crowdfunding success.

3.2.1 Identification of social ties from large social data dataset - Case: Indie Game Developer Community

For the first case study, Case: Indie Game Developer community, the goal was to build and demonstrate a new approach to identify relevant social ties from a large publicly available social media dataset. In order to fulfill this goal, a set of criteria for finding a suitable and adequately challenging social media-based large community were first devised. These selection criteria were based on theory-based sampling (Patton 2002), focusing specifically on “phenomenon of interest”. The selected dataset is well suited to study tie strengthening as a phenomenon in the context of, for example, knowledge work. The following selection criteria were considered while making the selection.

- The community must be large from the perspective of membership size, active members, amount of posts, and the number of discussions among the members. Due to this, the task becomes challenging while there is, seemingly, much noise in the form of low-frequency interactions that may not promote tie strengthening. It creates a challenge for filtering out relevant discussions and posts, most likely leading to increased tie strength.
- The community needs to have been active for several years so that ties have already been established and strengthened between its members over time.
- Ideally, the community will consist of simple problem-solving activities without needing further tie building, as well as more long-term and goal-oriented services such as capability building and longer-term learning that benefit from the building of relationships within the community and the strengthening of ties. This creates an additional challenge and needs to filter out irrelevant interactions from most relevant reciprocal interactions. It is unlikely that significant tie strengthening will occur in very loose communities, such as photo and video sharing communities.
- The community should not be very centralized, and top-down managed since the strengthening of ties requires frequent interactions that are reciprocal in nature (Blumstein and Kollock 1988; Granovetter 1973).
- The community should be open for access, have an open wall or API for collecting data, and be mainly or exclusively English. This will enable the use of content analysis and text mining algoritwasin later phases.

After examining different types of communities, both work- and non-work-related ones, as well as assessing their appropriateness and challenge for filtering relevant discussions from less relevant ones and taking into consideration the above-mentioned selection criteria, the Indie Game Developer community was selected.

3.2.2 Identification and role of different kinds of social ties using publicly available data about an event - Case: CMAD

The case study environment for Case: CMAD was community managers' online discussions in social media in connection to yearly- organized Community Manager Appreciation Day (CMAD2016) event that took place on January 25, 2016, in Jyväskylä Finland, and had 270 event participants. Case CMAD was selected because it satisfied the conditions suggested by Yin (1994) for a single-case design-based case study. The case CMAD was an extreme or unique case suitable to the overall goal of the research, which was to identify different types of social ties in a professional event using publicly available social media data. Following were the main reasons which show that CMAD was an appropriate case for addressing the research objective of identifying social ties from publicly available social media data.

- CMAD was a professional event with most participants belonging to the community managers' community. These participants can be viewed as advanced lead users of social media and online community management approaches. Most of them are highly active in social media (Aramo-Immonen, Jussila and Huhtamäki 2015; Aramo-Immonen, Kärkkäinen et al. 2016).
- As well as being active on social media in general, these event participants use social media in CMAD for purposes such as networking and maintaining relationships (see Aramo-Immonen, Jussila and Huhtamäki 2015; Aramo-Immonen, Kärkkäinen et al. 2016).
- Data related to the event CMAD are publicly available, which is essential for the main research problem addressed by this study.
- Based on prior studies of community managers in Finland (see Aramo-Immonen, Jussila and Huhtamäki 2015; Aramo-Immonen, Kärkkäinen et al. 2016), it can be demonstrated that community managers interact with each other also between events, and have also participated actively in planning the event. Hence, it can be assumed that gathering data based on these community members' discussions from Twitter and Facebook it can capture a sufficient and representative amount of data to draw conclusions.

Thus, the event CMAD 2016 was used as the second case for this dissertation to address this dissertation's research questions.

3.2.3 Identification and role of social ties in crowdfunding success - Experiment: Kickstarter

For the third sub-study, Experiment: Kickstarter, the goal was to identify the implicit social ties from social data and also to understand how these implicit social ties impact the outcome of the business phenomena or the business decision-making. In order to carry out this sub-study, the context of crowdfunding was selected. Crowdfunding has provided an alternative means to connect people that need funds for carrying out a project or an idea with people who are willing to fund those endeavors. (Madrazo-Lemarroy, Barajas-Portas and Labastida Tovar 2019) Existing research has found that social media can play an important role in enabling potential crowdfunding project owners to reach their project funding goals. Thus, this sub-study focused on identifying the implicit social ties from publicly available social data related to crowdfunding projects and the role of these implicit social ties towards crowdfunding project success.

The sub-study was further divided into two different parts. Separate hypotheses were developed for each of these parts. The objective of the first part of the sub-study was to understand the role of implicit social ties in general towards the success of a crowdfunding project. The aim of the second part of the sub-study was to understand the role of implicit social ties of project owners and project backers towards the degree of success of a crowdfunding project. Both parts of this sub-study were carried out using crowdfunding project data from the largest crowdfunding platform Kickstarter.com and the social media platform Twitter. Hence, the Experiment: Kickstarter, was used to address the research questions of this dissertation.

3.3 Data collection and analysis methods

In order to address the research questions of this dissertation three different studies were done. In order to carry out these studies different datasets were collected and different kind of data analysis were done. In the following subsections the details related to the dataset and analysis used in each of these study is presented.

3.3.1 Identification of social ties from large social data dataset - Case: Indie Games Developer Community

3.3.1.1 Data collection

Table 3.1 Details of Data Corpus Related to Content Attributes

Content Atkindsute	
Posts ,	26290
Comments	205729
Comment Replies	23084
Likes on Posts	332886
Likes on Comments	35152

In order to carry out this study, the social media data related to the indie games developer community was collected from Facebook. The Facebook data was collected using the Social Data Analytics Tool (SODATO) (Hussain and Vatrapu 2014; Hussain, Vatrapu et al. 2014). SODATO allows the systematic gathering, storing, and retrieval of social data across Facebook walls and groups. Using SODATO, the historical fetch of the Facebook group was conducted from 05-01-2008 to 29-04-2017. This study used data from 01-01-2015 to 31-12-2015. The details for this data corpus can be found in Table 3.1 and Table 3.2.

Table 3.2 Details of Data Corpus Related to Actor Attributes

Actor Attribute	Value
Total Unique Actors	337425
Unique Posters	24553
Unique Commenters	103581
Unique Comment Reply Actors	9613
Unique Wall Post Likers	327
Unique Comment Likers	16043
Unique Comment Reply Likers	5817

3.3.1.2 Data analysis methods

Social Network Creation: Facebook data allows for forthright analysis in general. The entities used to analyze this case study were Facebook posts, comments, and comment replies (replies to comments). To analyze the above-mentioned entities in Facebook data, a tailored Python script was written. This python script then transformed the refined data into a network. The network presents interconnections between people interacting on Facebook. The term "interconnections" refers to users initiating Facebook posts, comments, and replying to Facebook entities previously mentioned. Using the Network X library (version 1.11), the Python script built the network and serialized it in Graph Exchange XML Format or GEXF (version 1.2). The Python script created the network graph files.

Social Network Visualization: An open-source visualization and exploration platform, Gephi, (Bastian, Heymann and Jacomy 2009) was used to analyze and visualize the networks. In the analysis, Gephi was used to layout the networks, calculate metrics for the network nodes, filter out the network using different Gephi filters (such as self-loop, mutual edge, degree range), and adjust the visual properties of the visualized network. In this study, the layout of the networks was determined by a force-driven layout algorithm in which the nodes repel each other while the edges link them to act as springs to pull them back together (Blondel et al. 2008). As a result, the nodes that are interconnected will be located close together.

3.3.2 Identification and role of different kinds of social ties using publicly available data about event - Case: CMAD

3.3.2.1 Data collection

In order to carry out this study, two different sources of data were used. The first data source was the social media data from Twitter and Facebook related to the event CMAD2016. The second data source was a survey that was sent to the participants of CMAD2016. The details related to both the data sources are provided below.

Social Media Data: The social media data for CMAD2016 was collected from Facebook and Twitter.

Facebook Data: The Facebook data was collected using the Social Data Analytics Tool (SODATO) (Hussain and Vatrapu 2014; Hussain, Vatrapu et al. 2014). SODATO allows the systematic gathering, storing, and retrieval of social data across Facebook walls and groups. Using SODATO, the historical fetch of two Facebook pages (CMADFI 2014, CMADFI 2015) was conducted from 01-01-2014 to 26-05-2016 was. This study used data from 01-09-2015 to 30-04-2016. The details for the Facebook data corpus can be found in Table 3.3 and Table 3.4.

Table 3.3 Details of Data Corpus Related to Content Attributes

Content Attribute	Value
Posts	555
Comments	2925
Comment Replies	149
Likes on Posts	2529
Likes on Comments	2536
Likes on Comment Replies	104

Table 3.4 Details of Data Corpus Related to Actor Attributes

Actor Attribute	Value
Total Actors	8798
Total Unique Actors	374
Unique Posters	81
Unique Commenters	199
Unique Comment Reply Actors	53
Unique Wall Post Likers	327
Unique Comment Likers	204
Unique Comment Reply Likers	38

Twitter Data: Data on Twitter was collected in two phases.Flockler, a social media-driven content management system used to run the CMADFI website, was used to list all tweets sent before, during, and after CMADFI 2016. Flockler provides a web application programming interface (API) to gather all tweets associated with CMADFI 2016. Flockler, a social media-driven content management system used to run the CMADFI website, was used to list all tweets sent before, during, and after CMADFI 2016. Flockler provides a web application programming interface (API) to gather all tweets associated with CMADFI 2016. By processing the tweet ids from Flockler data, a complete set of Twitter data was collected, including the complete set of metadata provided to Twitter by each tweet. In order to access complete tweet data, including tweet sender, Twitter users mentioned in tweets, and hashtags, a batch script that utilizes Twitter REST API was implemented. The batch script exports tweet data in JSON for further processing. Using the batch script, tweets are exported from Twitter in JSON format for further processing. This study used data from 01-09-2015 to 30-04-2016. The details for the Twitter data corpus can be found in Table 3.5 and Table 3.6.

Table 3.5 Details of Data Corpus Related to Content Attributes

Content Attribute	Value
Total Tweets	555
Original Tweets	2925
Retweets	149

Table 3.6 Details of Data Corpus Related to Actor Attributes

Actor Attribute	Value
Total Users	12454
Total Unique Users	1651
Unique Original Tweet Users	858
Unique Retweet Users	1262

Survey Data: The second source of data was collected directly from conference participants. Self-reported data were essential in interpreting the social media data against the theoretical framing. Operationalizing the survey was based on the theoretical descriptions of Granovetter and the operationalized scale from Petroczi, Bazsó and Nepusz 2007. The following survey items were operationalized and are shown in Table 3.7. The survey was designed to capture the perceptions of event participants on their strong ties and possible weak ties from the event participants. Since weak ties are more numerous Granovetter 1973 and are, therefore, harder to recall by self-report, no questions related directly to them were included.

Table 3.7 Survey Questions

Q1	Which 3 - 5 CMAD 2016 participants do you interact most frequently with ?
Q2	Which 3 - 5 CMAD 2016 participants would you most likely ask a personal favor from or return personal favor ?
Q3	Which 3 - 5 CMAD 2016 participants have you known the longest in a professional context?
Q4	Which 3 -5 CMAD 2016 participants do you consider as your closest friend?
Q5	How novel (on average) was the information, you received from the CMAD 2016 participants amongst the following groups?
Q6	Which 3 - 5 CMAD 2016 participants do you consider as a source of most novel information or ideas?

The survey question 1 to 4 were designed to determine the strong ties of survey respondents. Since it was difficult to recall the names of all event participants, a maximum of five names was used. As part of question 5, participants rated the novelty of information derived from three groups of participants on a scale of 1-7. The three groups consisted of survey respondents, people they knew well, people they met face to face but did not know well, and people they had no face-to-face contact with. In general, Question 5 was used to identify the different sources and quality of the information. In Question 6, respondents were asked to identify novel sources of information for themselves. CMAD 2016 participants were provided with a link to an online survey via the CMAD Facebook group wall and the official CMAD Twitter handle. The survey was available in English and Finnish and was based only on the CMAD 2016 event. Twenty-five survey responses were received from a total of 270 participants.

3.3.2.2 Data analysis methods

Temporal Analysis: A temporal analysis of Twitter and Facebook data was conducted using data warehousing and online analytical processing (OLAP) technology using Microsoft SQL Server. Data from Twitter and Facebook was analyzed using a multi-dimensional data model utilizing interactions as numeric measures. Further processing of interaction measure data took into account multiple dimensions: temporal (daily, weekly, monthly, and yearly), actions (post, comment, like), and artifacts (posts, comments, tweets, and retweets).

Data Processing in Social Networks: The data from Twitter and Facebook can generally be analyzed straightforwardly. In the case of Twitter, the used REST API makes the tweet data easy to process programmatically. In the case of Twitter, users (e.g., @menonkaran) and hashtags (e.g., cmadfi) are represented by specific syntax and structure. In the analysis of Facebook, the components for the analysis were posts, comments, replies, and likes. Using a Python script, the entities mentioned above were identified in both Twitter and Facebook data. After the data were refined, the script transformed the data into two networks:

- In the first network, people are connected by conversations over Twitter. In particular, users mentioning one another in tweets via comments and discussions are interconnected.
- The second network shows interconnections between people communicating on Facebook. More specifically, interconnections refer to users initiating Facebook posts, comments, comment replies, and "Likes" to aforementioned Facebook entities.

For the construction of the network, the Python script uses the NetworkX library (version 1.11). It serializes it in Graph Exchange XML Format (version 1.2).

Social Network Analysis: The networks were analyzed and visualized using Gephi, an open-source tool for interactive visual exploration and visualization (Bastian, Heymann and Jacomy 2009). The layout of networks, calculations of metrics for network nodes, and analysis of subnetworks (e.g., egocentric networks of individual nodes) were carried out using Gephi. Additionally, it was used to identify clusters (Modularity Class metrics) through Gephi's implementation of the community detection algorithm (Blondel et al. 2008) and adjust its visual properties according to the analysis. The text related to identified different modularity classes was read and analyzed. This was used to identify the related theme for each modularity class. Communication frequency was used as a proxy for tie strength, in this case, to evaluate tie strength at the interpersonal level (between the event participants). In the analysis, two metrics were of interest, weighted degree (sum of weighted in-degree and out-degree) and modularity class (clustering). These networks were designed using a force-driven layout algorithm in which nodes repel each other, while the edges connecting the nodes act like springs that pull the nodes back together (Blondel et al. 2008). As a result, the nodes that interact will be near each other.

3.3.3 Identification and role of social ties in crowdfunding success - Experiment: Kickstarter

3.3.3.1 Data collection

The study was conducted using two different datasets. The first dataset was collected from a crowdfunding platform. The second dataset was conducted from a social media platform.

Crowdfunding Platform Data: The crowdfunding data for this study was collected from the crowdfunding platform Kickstarter.com. Kickstarter is one of the largest online crowdfunding platforms globally that enables businesses or teams to issue funds over the internet and receive small investments from registered funders in return. Kickstarter.com is a reward-based crowdfunding platform that follows the AON model. As of Sep 2021, Kickstarter has successfully funded 208,212 projects with a total pledge funding of 6.10 billion US dollars with an overall success rate of 39.089%. For this study, Kickstarter project data were collected from webrobots.io, which uses a scraper robot to crawl Kickstarter project data and provide it in both .csv and .json formats. Webrobots.io crawls the data monthly, and the complete dataset is accessible through a link on the website. The initial dataset used in this study includes data on all Kickstarter projects between April 2009 and September

2018. Only the crowdfunding projects conducted between June 2016 and September 2018 were selected from this dataset. Projects whose related tweets were unavailable back to six months before the project's start date were filtered out of these crowdfunding projects. There were a total of 2,161 crowdfunding projects that satisfied all of the criteria used to filter the data. The crowdfunding project data contains attributes such as campaign organizer IDs, project creator name, webpage URLs, shortened versions of these webpage URLs, campaign fundraising goals, fundraising durations, funding amounts raised, campaign start dates, campaign end dates, and projects' country of origin.

Social Media Data: The social media data used in this study was collected from the social media platform Twitter. Twitter data were collected by writing a search query containing keywords related to the Kickstarter platform and using the Twitter Premium API. This query contained terms such as "#Kickstarter," "#kickstarter," "@kickstarter," and related terms. Each tweet was recorded with its timestamp, author handle, any user handles mentioned, the text, and any URLs that appeared in the tweet's text or user metadata. Data from Twitter were collected from June 2016 to September 2018. A detailed description of the Twitter dataset collected is found in Table 3.8 and the filetered dataset is found in Table 3.8.

Table 3.8 Data description of the collected Twitter dataset

Variables	Value
Total number of tweets	4,206,408
Total number of retweets	1,856,063
Total number of users	4,406,408
Total number of unique users	1,128,397

Table 3.9 Data description of the filtered Twitter dataset

Variables	Value
Total number of tweets	172,892
Total number of retweets	26,350
Total number of users	255,419
Total number of unique users	21,491

3.3.3.2 Data analysis methods:

Data processing for Implicit Network Creation: In order to perform the data analysis, the data had to be cleaned and organized. Based on the Python code script that was used in the data cleaning process, a total of 172,892 tweets from 2,161 Kickstarter projects were identified and used for this study. Each Kickstarter project's tweets were filtered, only using tweets posted during the crowdfunding project campaign for that Kickstarter project. For each Kickstarter project, a separate network was created through the tweets of the respective project. The network represents interconnections between people interacting over Twitter. Interconnections explicitly point to users mentioning each other in tweets. These separate implicit social networks for each Kickstarter project were used to calculate various network metrics.

Regression Analysis: Regression analysis is a set of statistical processes for estimating the relationships between a dependent variable (also known as 'outcome variable') and one or more independent variables (also known as 'predictors', 'covariates', or 'features') (Freedman 2009). In this study, the experiment was divided into two parts and two different sets of hypothesis were developed. The first part of the experiment was related to understanding if there is any association between implicit social ties and crowdfunding success. For this part of the experiment the dependent variable was binary: there were two possible outcomes, either the crowdfunding project achieves its funding goal (success) or the crowdfunding project does not achieve its funding goal(fail).The logistic regression model is the most frequently used regression model for this kind of discrete data (Hosmer, Lemeshow and Sturdivant 2013), and, hence, it was used. The second part of the experiment was related to understanding the role of the implicit social ties of the project owners and potential project backers on the degree of crowdfunding project success. In this case, linear regression was used to study the relationship between the degree of crowdfunding project success (dependent variable) and the implicit social ties of the crowdfunding project owner and potential project backers (independent variables).

3.3.3.3 Hypothesis development

This experiment can be further divided into two different parts. The first part of the experiment henceforth referred to as *Exp: Role of implicit social ties in crowdfunding* *project success* was used to understand the role of implicit social ties in general towards the success of the crowdfunding project. In order to understand this, a set of hypotheses were developed and tested. The second part of the experiment henceforth referred to as *Exp: Role of implicit social ties of project owners and project backers in the degree of success of crowdfunding project* was used to understand the role of implicit social ties of project owners and potential project backers towards the degree of success of crowdfunding project. In order to understand this, a different set of hypotheses were developed and tested. The logic behind the development of the hypothesis related to Exp: Role of implicit social ties in crowdfunding success and Exp: Role of implicit social ties of project owners and project backers in the degree of success of crowdfunding project is described below.

Exp: Role of implicit social ties in crowdfunding project success The goal of this part of the experiment was to investigate the four dimensions of tie strength to examine the role of implicit ties captured in social media on crowdfunding projects' success. These four dimensions of tie strength are listed below:

- Emotional intensity,
- Intimacy,
- Reciprocal services and
- Structural dimensions

Data from Twitter was used to develop seven different measures measuring the different dimensions of tie strength regarding implicit ties. Metrics such as the number of tweets and the number of retweets are based on the conversation interaction between the crowdfunding project users on Twitter. The average number of followers per potential project backer and average retweets per potential project backer can be determined by social media's aggregated relationship data and user profile data collected from social media. Textual data from social media is used to derive the measure, such as the number of mentions. A combination of conversation and textual data was used to create a social network based on the frequency of interaction. Degree centrality and betweenness centrality were calculated for these networks.

The emotional intensity dimension of tie strength measures the overall strength of the emotional ties between two entities rather than a specific direction of positive or negative feelings. People with great intense relationships will spend more time together than people with less intense relationships and have a higher tie strength (Granovetter 1973; Mathews et al. 1998). Earlier research has confirmed that the tie strength's emotional intensity dimension influences buyers' commitment to the selling organization (Stanko, Bonner and Calantone 2007). This has been found to support the buyer-seller relationships and the social pull towards agreement (Gilliland and Bello 2002).

Furthermore, research studies by Gilbert 2012; Kumar et al. 2013; Luarn and Chiu 2015; Pir Mohammadiani, Mohammadi and Malik 2017 have proposed measuring the emotional intensity dimension of tie strength in social media by measures like the number of tweets. Thus, the following hypothesis related to the emotional dimension of tie strength and success of the crowdfunding project is proposed:

X1: *A positive association exists between the emotional intensity dimension of implicit ties on social media and crowdfunding success.*

The intimacy dimension of tie strength is defined as a sense of reliance and security provided by deep affection between two entities (Marsden and Campbell 1984; Petroczi, Bazsó and Nepusz 2007). According to Stanko, Bonner and Calantone 2007, tie strength's intimacy dimension affects consumers' commitment to a business organization. Based on prior studies related to social media-based tie strength by Gilbert 2012; Kumar et al. 2013; Pir Mohammadiani, Mohammadi and Malik 2017 , measures like the number of mentions and the average number of followers per user have been applied in research to evaluate the intimacy dimension of tie strength. Based on the above ground, it is reasonable to use the number of mentions and the average number of followers per user to measure the tie strength's intimacy dimension. The greater the intimacy of tie strength in the social media network for a crowdfunding project, the more likely the crowdfunding project will succeed. Thus the following hypothesis is proposed :

X2: *A positive association exists between the initimacy dimension of implicit ties on social media and crowdfunding success.*

Reciprocal service dimension refers to the various forms of interaction or specific services or those related to communication. Some of the significant studies related to tie strength suggested that strong ties are motivated to share information, knowledge, or resources, thus providing ready access to information circulating in their network (Granovetter 1973; Krackhardt 1992; Levin and Cross 2004). Reciprocal services refer to the extent to which the parties take responsibility for the partner firm's well-being as well as their own. The firms can demonstrate a willingness to do things for each other for the good of the relationship. Both firms place great value on their relationships, and there is a unity of interest and solidarity (Heide and John 1992; Noordewier, John and Nevin 1990).

Social media data measures such as the number of retweets and the average number of followers per user were previously used to evaluate reciprocal services of tie strength (Gilbert 2012; Gupta, Menon et al. 2016; Kumar et al. 2013; Pir Mohammadiani, Mohammadi and Malik 2017). Following the earlier research, measuring the reciprocal services dimension of tie strength with the number of retweets and the average number of retweets per user is reasonable. Also, it can be assumed that the reciprocal services dimension of tie strength of a crowdfunding project in social media will positively impact the success of the crowdfunding project. Hence, the following hypothesis is proposed:

X3: *A positive association exists between the reciprocal services dimension of implicit ties on social media and crowdfunding success.*

According to the structural dimension of tie strength, the ties are related to social homogeneity, shared affiliation, overlap among social circles, the topology of the network, and informal social networks (Alba and Kadushin 1976; Burt 2004; d. m. b. d. m. and Ellison 2007; Xiang, Neville and Rogati 2010). Some past studies suggest that strong ties tend to bond similar people together, and these similar people tend to cluster together such that they are all mutually connected (Burt 2004; Gilbert 2012; N. Lin, Dayton and Greenwald 1978). In previous studies on social media-based tie strength by Boyd, Golder and Lotan 2010; Gilbert 2012; Hutto, Yardi and Gilbert 2013, degree centrality and betweenness centrality were recommended as two important factors to evaluate the structural dimension of tie strength.

According to prior literature, structural dimensions of tie strength in social networks can lead to positive outcomes, such as project success and relationship development. This study uses degree centrality and betweenness centrality to measure the structural dimension of tie strength of implicit ties in social media, following previous research. Based on the earlier research, it is assumed that the structural dimension is positively associated with the crowdfunding project's success. Thus, the following hypothesis is proposed:

X4: *A positive association exists between the structural dimension of implicit ties on social media and crowdfunding success.*

Exp: Role of implicit social ties of project owners and project backers in degree of success of crowdfunding project The goal of this part of the experiment was to investigate the role of implicit social ties of project owner's and potential project backer's towards the degree of success of crowdfunding project.

Crowdfunding success can be defined in several ways, such as binary (that is, the project is either successful or unsuccessful) (Madrazo-Lemarroy, Barajas-Portas and Labastida Tovar 2019; Mollick 2014) or as the ratio between a project's funding goal and its actual funding received by a project (Shneor and Vik 2020). In this study, crowdfunding success is defined as the ratio of the funding a project has received to its funding goal. This has been argued to provide more comprehensive insights into crowdfunding project success than just using a binary value for crowdfunding project success (C. Yin, L. Liu and Mirkovski 2019). Based on the current study's use of Twitter data to measure social ties, four measures are included in the model to reflect social ties on Twitter: retweet counts, followers, degree centrality, and betweenness centrality.

Two factors related to network structure are degree centrality and betweenness centrality, which can be computed based on a social network and applied to identify different types of social ties (Borgatti and Halgin 2011; Marsden and Campbell 2012). In many other studies of networks, explicit relationship-related social media data (such as Facebook friends) has been used to construct social networks. However, this approach may be hard to apply in the crowdfunding context (Y. H. Tan and Reddy 2021). Thus, in this study, an implicit social network based on the interaction between potential project backers and project owners vis-à-vis a crowdfunding project on Twitter. In more explicit terms, the user's mentioning of each other it tweets is considered the interaction. Network nodes represent project owners and potential backers, while the network edges reflect the number of interactions between these different nodes. A similar approach has been used to study a variety of phenomena in the past(Aramo-Immonen, Jussila and Huhtamäki 2015; Aramo-Immonen, Kärkkäinen et al. 2016).

One commonly used measure derived from Twitter data is retweet counts in the research related to social ties. With the use of social media, a new kind of user behavior has emerged - that is, the use of a friend's activity to distribute a duplicate of the content they post online (words, videos, or pictures) (Geva, Oestreicher-Singer and Saar-Tsechansky 2019). Retweeting is Twitter's version of this behavior. A lot of studies on the social ties involved in studying different phenomena have analyzed retweets.(Geva, Oestreicher-Singer and Saar-Tsechansky 2019; Zhan Shi, Huaxia Rui and Whinston 2014) Few studies have noted that retweet functionality positively affects the success of crowdfunding campaigns (J. Liu and Ding 2020). Project owners and project backers can both distribute crowdfunding content on Twitter. Thus, it is assumed that the retweet counts of both project owners and potential project backers affect degrees of crowdfunding success. Hence, we propose the following hypotheses:

Y1a: *Project owners' retweet counts in a crowdfunding project are positively associated with the crowdfunding project's degree of success.*

Y1b: *Potential project backers' (average) retweet counts in a crowdfunding project are positively associated with the crowdfunding project's degree of success.*

Following functionality is another variable commonly used to measure social ties on Twitter. Usually, in these studies, explicit relationship data (follower - followee user list) related to who follows whom on Twitter was used to create social networks and calculate the number of followers of the user (Gilbert 2012; Zhan Shi, Huaxia Rui and Whinston 2014). Hence, these studies used data associated with explicit ties in social media. However, this study used Twitter user metadata to get an aggregate measure of a user's total followers. Even so, this measure does not include explicit data about relationships (follower - followee lists) and cannot be used to build social networks. Accordingly, in this study, the number of followers refers to a measure of implicit ties instead of explicit ties, as in earlier studies.

Many studies have used this measure in the crowdfunding context but predominantly only project owners' number of followers Hong, Hu and Burtch 2018; C. T. Lu et al. 2014. In many of these studies, it has been found that followers positively affect crowdfunding success Kang, Jiang and C. H. Tan 2017; Kim and Z. Zhang 2018; Mollick 2014. The number of followers provides a preliminary indicator of the size of the social network among project owners and potential backers. Accordingly, this study hypothesizes that followers of both project owners and potential project backers affect the success of crowdfunding projects. The following hypotheses are proposed:

Y2a: *Project owners' number of followers in a crowdfunding project is positively associated with the crowdfunding project's degree of success.*

Y2b: *Potential project backers' (average) follower counts in a crowdfunding project* *are positively associated with the crowdfunding project's degree of success.*

According to the research on social networks and the research related to identifying social ties, nodes' relational positions play a crucial role in terms of their social ties (Borgatti and Halgin 2011). Several metrics measure and capture the relational properties of nodes. These metrics are commonly known as centrality metrics (Chen et al. 2012; Freeman 1978). Based on graph theory, degree centrality is defined as the number of neighbors a node has. By analyzing the number of links held by each node, degree centrality identifies individuals who possess the most information or who can connect quickly with others. In contrast, betweenness centrality measures centrality in a graph-based on shortest paths, and it measures the number of times a node lies on the shortest path. It can be used to analyze communication dynamics in a network, and it can be used to identify individuals who influence the information flow in a network (Hansen et al. 2020). Therefore, the degree centrality and betweenness centrality of project owners and potential project backers on social media can influence a crowdfunding project's information flow and reach. Recent studies by (Chung, Y. Li and Jia 2021) and (Y. H. Tan and Reddy 2021) found that project backers' social networks' degree centrality and betweenness centrality positively affect backers' pledge decisions regarding crowdfunding projects. Therefore, it can be reasonably assumed that the degree centrality and betweenness centrality of the owners and potential backers of a crowdfunding project can influence their success. Thus, the following hypotheses related to the network measures of project owners and potential project backers is proposed:

Y3a: *Project owners' degree centrality in a crowdfunding project is positively associated with the crowdfunding project's degree of success.*

Y3b: *Potential project backers' (average) degree centrality in a crowdfunding project is positively associated with the crowdfunding project's degree of success.*

Y4a: *Project owners' betweenness centrality in a crowdfunding project is positively associated with the crowdfunding project's degree of success.*

Y4b: *Potential project backers' (average) betweenness centrality in a crowdfunding project is positively associated with the crowdfunding project's degree of success.*

3.3.3.4 Data measures

In this subsection the different data measures which were used in the two parts of the experiment are described. First the data measures related to *Exp: Role of implicit* *social ties in crowdfunding success* are described. Second the data measures related to *Exp: Role of implicit social ties of project owners and project backers in degree of success of crowdfunding project* are described.

Exp: Role of implicit social ties in crowdfunding project success The different data measures used are described below:

- **Number of Mentions :** On Twitter, a tweet can contain another user's username: this is known as a mention. Tweets can contain multiple usernames, and each username is added to the text of the tweet by adding the @ symbol before it. Example: "I'm tweeting about @user1@user2!" User1 and user2 are both Twitter mentions. In this study, "number of mentions" refers to how many mentions each Kickstarter project received from all its tweets.
- **Number of Tweets :** In this case, it refers to the total number of tweets about each crowdfunding project during its campaign.
- **Number of Retweets :** It is possible for a Twitter user to publicly share a Tweet with others who follow them, known as retweeting. Twitter provides an identifier in the data to tell whether a tweet has been retweeted. In this study, the number of retweets refers to the total number of retweets found in all the different tweets from each Kickstarter project.
- **Average Retweet Count per Potential Backer :** Twitter also provides information about how many times each user has been retweeted. In this study, the average retweet count per potential backer refers to the average number of retweets of each user present in the individual Kickstarter project's different tweets.
- **Average Number of Followers per Potential Backer :** On Twitter, a user can follow another user as a follower. Followers receive tweets from the users they follow, and Twitter shows the number of followers a Twitter user has. In this study, the average number of followers per potential backer refers to the average number of followers of each user present in the different tweets of the individual Kickstarter project.
- **Degree Centrality of Project Owner :** Graph theory defines degree centrality as the number of neighbors a given node has, and degree centrality

counts links held by each node. Degree centrality is useful for analyzing individuals who possess the most information or can quickly connect to the broader network. (Hansen et al. 2020) A separate network was created from each Kickstarter project's tweets, representing the interconnections between people communicating over Twitter. In more explicit terms, it refers to users mentioning each other in tweets. In this study, the degree centrality refers to the degree of centrality of the Kickstarter project owner in the Twitter network created from the project's tweets.

- **Betweenness Centrality of Project Owner :** In graph theory, betweenness centrality is a measure of centrality in a graph, and it measures the number of times a node lies on the shortest path between other nodes. Analyzing betweenness is helpful in understanding communication dynamics in a network and identifying the individuals who influence its flow. This study created a separate network for each Kickstarter project from each Kickstarter project's tweets. The network represents interconnections between people communicating over Twitter. In more explicit terms, it refers to users mentioning each other in tweets. In this study, the degree centrality refers to the degree centrality of the betweenness project owner in the Twitter network created from the project's tweets.
- **Project success :** The project success is coded as a binary variable. Its value is zero if the project funding goal is not achieved (amount of project funding less than project fund goal), and its value is one when the project funding goal is achieved (amount of project funding equal to or more than project fund goal).

Exp: Role of implicit social ties of project owners and project backers in degree of success of crowdfunding project In this study, four measures of project owners' and potential project backers' social ties, respectively, were set as independent variables supposed to be associated with a crowdfunding project's degree of success. For example, a crowdfunding project that reached 50% of its project funding goal would have a degree of success equal to 0.5; meanwhile, a crowdfunding project that reached 100% of its project funding goal would have a degree of success equal to 1. The different data measures used are described below:

- **Retweet count of project owners :** It refers to a project owner's number of retweets in a crowdfunding project.

- **Number of followers of project owners :** It refers to a project owner's number of followers in a crowdfunding project.
- **Betweenness centrality of project owners :** It refers to a project owner's betweenness centrality, calculated from the implicit social network created for each crowdfunding project using tweets related to the project.
- **Degree centrality of project owners :** It refers to a project owner's degree centrality, calculated from the implicit social network created for each crowdfunding project using the tweets related to the project.
- **Retweet count of potential project backers :** It refers to a potential project backer' s average number of retweets in a crowdfunding project.
- **Number of followers of potential project backers :** It refers to a potential project backer's average number of followers in a crowdfunding project.
- **Betweenness centrality of potential project backers :** It refers to the average betweenness centrality of a potential project backer, calculated from the implicit social network created for each crowdfunding project using tweets related to the project.
- **Degree centrality of potential project backers :** It refers to the average degree centrality of a potential project backer, calculated from the implicit social network created for each crowdfunding project using tweets related to the project.
- **Degree of success of crowdfunding project :** It refers to the ratio of the project funding received (in US dollars) to the project funding goal (in US dollars).

3.4 Brief summary of the different datasets and methods used in the dissertation

Based on the description provided about the different datasets used in this dissertation, it can be observed that the Case: CMAD used a relatively small size dataset. On the other hand, the datasets used in Case: Indie Game Developer Community and Experiment: Kickstarter used large-size datasets. Table 3.10 provides a brief description of the different data sets used in this dissertation. The table also presents the different research papers where partial results of the different sub-studies of this dissertation were presented.

This section also provides details related to the role played by the author of this dissertation in carrying out the various sub-studies. It also explains the role and tasks which were performed by me in order to complete the sub-studies. The partial results of the different sub-studies were presented in the form of different research papers as highlighted in Table 3.10. For the sub-study Case: Indie Game Developer Community, I developed the initial idea of the study with other co-researchers. I carried out the data preparation and data analysis. I was primarily responsible for developing the method related to building the reciprocal interaction network and also doing the analysis of the data.

For the sub-study Case: CMAD, I developed the initial idea of the study with other co-researchers. I identified the relevant literature and the theoretical background and concepts needed for carrying out this study. I carried out the initial data preparation and data analysis. I developed the initial methodology for carrying out the sub-study with the help of other co-researchers. The findings of the sub-study and the conclusions which could be drawn from the results of the sub-study were developed with the guidance of other co-researcher.

For the third sub-study Experiment: Kickstarter, I developed the initial idea for research with other co-researchers. I identified the related literature and carried out the initial data preparation and data analysis. With the guidance of the other co-researchers I developed the different hypothesis for the sub-study and performed interpretation of the study results and conclusions of these results.

Table 3.10 Data Sources and Methods

Study	Case: Indie Game Developer Community	Case: CMAD	Experiment: Kickstarter
Research Question Addressed	RQ 1	RQ 2	RQ 3 , hypotheses an
Methodology and Strategy	Case Study	Case Study, Survey	Experiment
Methods	Social Network Analysis	Social Network Analysis, Descriptive data analysis, Temporal data analysis, Cluster analysis, Content analysis	Social Network Analysis, Logistic Regression, Linear Regression
Data	Facebook data about an Indie Game Developer Community	Social media data related to CMAD 2016 from Facebook and Twitter. Survey related to CMAD 2016	Crowdfunding platform data about crowdfunding projects from Kickstarter.com. Social media data from Twitter related to these crowdfunding projects.
Partial Results presented in Research Papers	Torro et al. 2017	Gupta, Menon et al. 2016 Gupta, Kärkkäinen, Torro et al. 2019	Gupta, Kärkkäinen, H. Li et al. 2022 Gupta, H. Li et al. 2022

4 RESULTS

4.1 Case: Indie Game Developer Community

This study developed an approach for filtering out the relevant social ties from a large publicly available social media data. The Python script used by this study to create the network graph resulted in a network with 16260 nodes and 123137 edges. The graph file (.gexf) was opened in Gephi to read the graph components and filter the network. First, the 'Giant Component' Gephi filter was applied. This filter filters out all the graph components other than the component that have the most nodes. Second, the Gephi filter 'Mutual Edge' was applied to the network. Mutual Edge filters only keep edges that have mutual or reciprocal edges. This step is essential as the tie-strengthening process relies on reciprocal interaction between actors (nodes). Therefore, all actors with no reciprocal interaction were filtered out. Third, the Self Loop filter was applied to the network. This filter eliminates all self-loops from the network. This step is essential since the tie strength is always between two dyads, not an individual node. Last, the 'Degree Range' Gephi filter was used, which keeps nodes that fall within the defined range.

Visualization of the filtered network was performed using force-driven layout algorithms. Figure 4.1 depicts the visualization of the Facebook group Indie Game Developers during the study. Figure 4.1 shows the filtered network resulting from applying the Gephi degrees range filter as explained above. The degree range for the degree range filter was set between 100 and 645, which yielded the filtered network of 313 nodes and 1072 edges. In Figure 4.1, the "nodes" represent the members of the Indie Games group. Connections between members indicate how the members interact with each other, but the thicker the connection, the stronger the interaction (line width). In this network, the node color represents a cluster of nodes based on a community-detection algorithm that examines the network and finds a cluster of nodes that are particularly closely interconnected. The size of the nodes represents the amount of interaction (sum of indegree and outdegree) between members. The larger the node size, the more interaction there is among members.

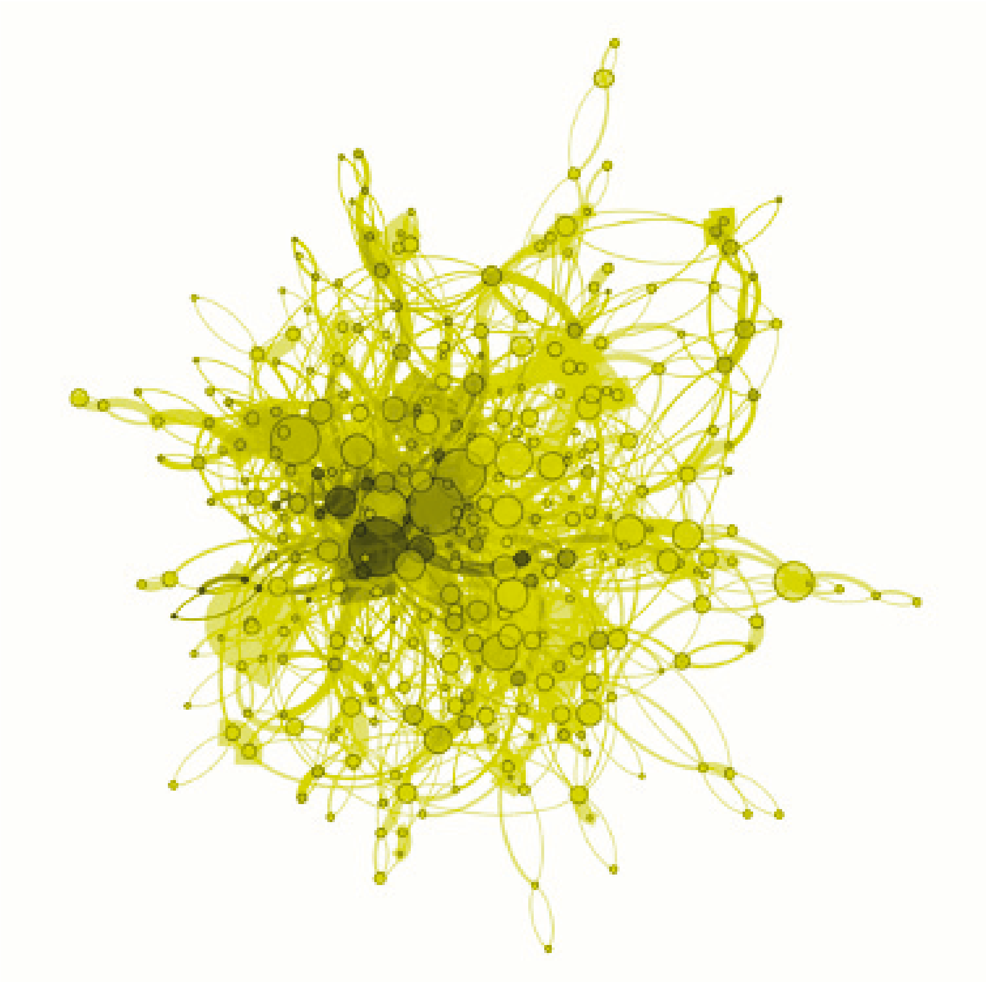

Figure 4.1 Force-driven network of people based on Facebook conversations.

As shown in Figure 4.1, these 313 nodes and 1072 edges illustrate the number of people (nodes) and their communication patterns (edges). The visualization has been cleared of noise, which means that fewer than 100 messages exchanged between individuals were not taken into account. Granularity may be adjusted as desired, of course. Noise removal is directed at low-frequency interactions (edges with low weight), in which strengthening ties is unlikely since the overall amount of communication remains low. Therefore, different tie strengthening patterns are expected to manifest within these 1072 edges. The same approach could also be applied to much larger data sets, leading to significant generalization capabilities when analyzing the content of these messages.

4.2 Case: CMAD

In this section, the findings of the different analysis which were performed in this study are presented.

4.2.1 Descriptive analysis

Literature has indicated that Twitter usage contributes to building new relationships. This was also true when examining Twitter data for CMAD2016. Below are some examples of CMAD2016 tweets used for information sharing, developing new relationships, and strengthening existing ones. The tweets were originally written in Finnish and have been translated.

- "Today Jyväskylä, some and #cmadfi. If you have networked communication and WordPress in mind, please contact melink ". This is an example of a tweet to establish a new tie.
- "Have a great #cmadfi-day in Jyväskylä These ladies won't be able to make it today in person, but we will be there in spirit and will follow live tweets! #yhteisömanagerit".

The questionnaire was completed by 24 of the 270 participants of CMAD2016, including 16 females and eight males belonging to different cities in Finland. Both the organizers (who were also participants) and event participants completed the questionnaire. Facebook conversations involved 30 participants who had ten or more conversations, 49 participants with five or more conversations, and 91 participants with at least one. In total, 119 Twitter participants had ten or more conversations, 134 participants had five or more conversations, and 214 participants at least had one conversation. As for question 5 in Table 3.7 related to the most novel information received by the questionnaire respondents, the average rating (on a scale of 1 to 7) for the three different options were: 5.13 for had not met face to face; 4.65 for had met face to face but did not know well; and 4.00 for knew well.

4.2.2 Temporal analysis of social media

CMAD2016's social media activity was monitored on Facebook and Twitter from 1st September 2015 to 30th April 2016. Figure 4.2 shows that there is only one significant spike in tweets and retweets during the week of the actual CMAD2016 event.

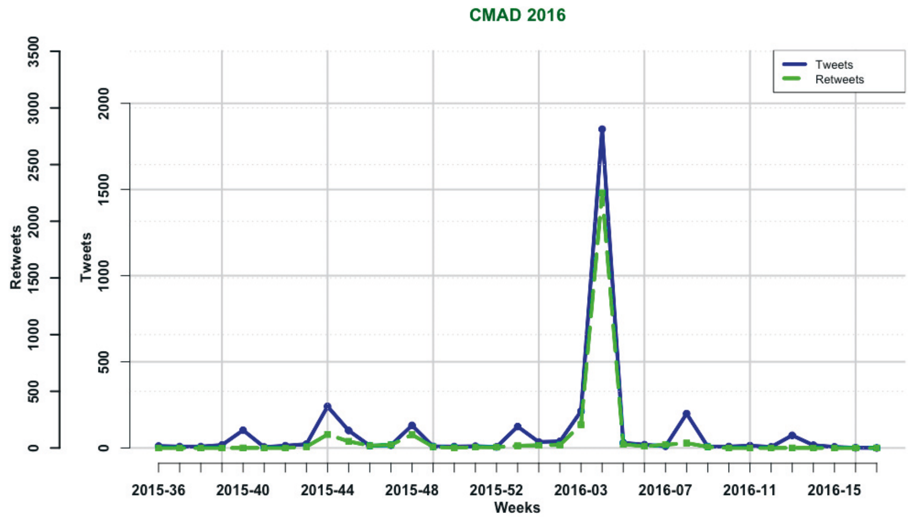

Figure 4.2 Temporal analysis from Twitter

Figure 4.3 shows that the number of comments, posts, and likes to the CMAD2016 Facebook page spiked during the weeks leading up to the event.

4.2.3 Analysis based on correlating social media data with survey data

Figures 4.4 and 4.5 show a visualization of the CMAD participants' conversations on Twitter and Facebook during the study period. The nodes in the visualization represent the CMAD participants. Furthermore, each participant's interests are made visible by their connections with other participants. The greater the interest, the greater the connection width (line width in Figures 4.4 and 4.5). A node's color indicates a cluster of nodes in the network, based on a community detection algorithm that analyzes the network to identify nodes that are particularly tightly interconnected.

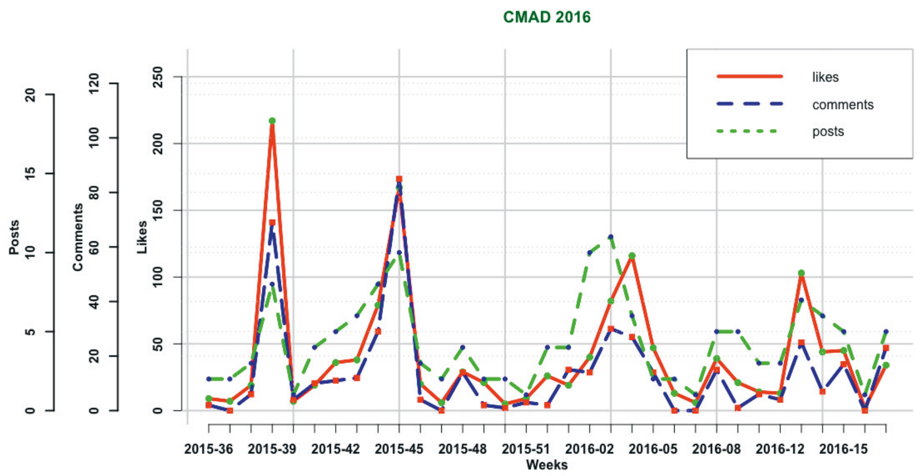

Figure 4.3 Temporal analysis from Facebook

The nodes labeled in the network graphs in Figures 4.4 and 4.5 represent the survey participants (alphabetical letters A to X) and their novel source of information as provided in the survey response for Question 6. For example, a survey participant is labeled as A while his or her novel sources of information are labeled A1, A2, A3, A4, and A5. In the case of Facebook (shown in Figure 4.5), four distinct clusters were found, whereas, in the case of Twitter (shown in Figure 4.4), 25 distinct clusters were found.

Content analysis of the identified clusters was performed. Following this content analysis (by reading tweets and Facebook posts), different potential sub-communities were identified (based on themes identified from the content analysis), and different clusters were labeled. As shown in Table 4.1, it was possible to identify these potential sub-communities from Twitter data.

Content analysis of Twitter data produced a number of possible sub-communities based on the distinct clusters (modularity class) identified from the conversational data. Modularity class 20, for example, had to do with the knowledge management discussion theme, while modularity class 1 dealt with personal branding. In contrast, the content analysis of the Facebook data did not reveal any specific themes. It could not be used to identify any sub-communities based on the different identified clusters.

In Question 6, 15 responses were received from the 25 survey participants regard-

Figure 4.4 Force driven network of people based on tweets

ing the participants whom a survey respondent found to be the most novel source of information. These responses were used to create Table 4.2, which shows the calculated modularity class (from social media data) of the survey respondent and the most novel information sources identified by each respondent. As shown in Table 4.2, column "Survey respondents" refers to 15 survey participants, which were coded by alphabetical letters; Columns I-V are the novel information sources that respondents identified. The "Modularity class number" pertains to the various modularity class-based clusters of survey respondents identified during the analysis. It can be seen in Table 4.2 for the Twitter data. Table 4.2 shows novel sources in different

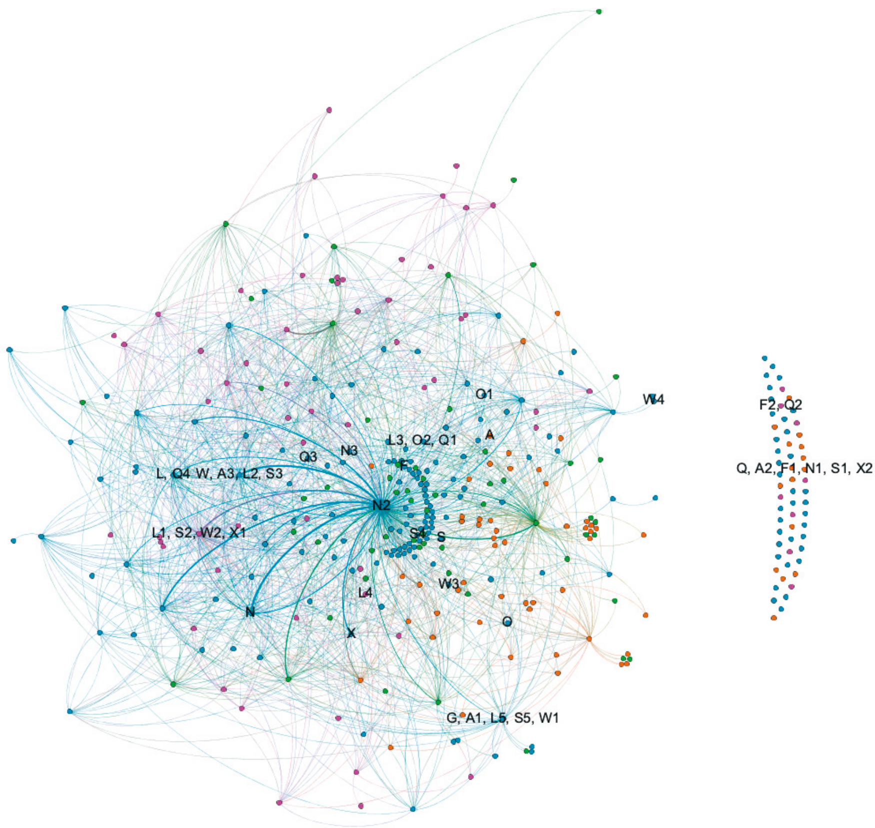

Figure 4.5 Force driven network of people based on Facebook data

modularity classes compared to survey respondents using the blue color.

Table 4.2 shows that the 15 survey respondents provided 55 unique information sources that correspond to the individual cells of the table. As shown in the table, 44 out of 55 novel information sources (approximately 80%) in the case of Twitter belong to a different modularity class. In contrast, only seven out of a total of 55 individual novel information sources on Facebook (approximately 12%) belong to a different modularity class. Furthermore, Table 4.2 shows that Twitter data contains all 55 individual novel sources. In comparison, Facebook data contains only 31 of the 55 individual novel sources.

Table 4.1 Identification of clusters based on modularity class using Twitter data

Cluster Name	Modularity Class
Personal branding	1
Employee advocacy	2
Drawings and infographics	3
**Broadly about cmadfi event	5
Community manager	6
Communications	7
*Reporting on CMADFI event	10
Customer service	15
Project	16
*Outsider greetings	17
TEKES	18
Knowledge management	20
Jyväskylän energia	21

Table 4.2 Correlating modularity class of novel information sources with social media data

Survey Respondents	Modularity Class of Respondents		Novel Source of Information & their Modularity Class↓									
			I		II		III		IV		V	
	T	F	T	F	T	F	T	F	T	F	T	F
L	20	1	20	0	16	1	1	1	7	0	16	1
R	7	-	5	-	16	-	2	-	5	-	20	-
Q	5	1	1	1	7	0	6	1	20	1	--	--
S	16	1	5	1	20	0	16	1	2	1	16	1
O	20	1	7	1	16	1	1	-	1	-	-	-
W	16	1	16	1	20	0	0	1	7	1	-	-
A	6	1	16	1	6	1	6	1	5	-	16	-
M	1	-	5	-	5	-	20	-	--	--	--	--
N	16	1	5	1	5	1	20	1	--	--	--	--
U	16	-	2	-	1	-	20	-	--	--	--	--
V	2	-	20	-	7	-	2	-	5	-	--	--
X	5	1	20	0	5	1	2	-	--	--	--	--
F	7	1	5	1	7	0	--	--	--	--	--	--
G	16	1	7	-	16	-	3	-	--	--	--	--
D	16	-	16	-	5	-	--	--	--	--	--	--

	Different Modularity Class than respondent
--	Did not provide a response
-	Data not available
T	Twitter data
F	Facebook data

4.2.4 Findings based on survey data

Based on Figure 4.5, it is evident that Facebook had one central node from which most of the connections were made. In addition, there were a few nodes in Facebook data that were not at all connected. These were people who posted something on the CMAD2016 Facebook page but did not receive a response. In Twitter (Fig. 4.4), there was no central node connected to all the other nodes.

For each survey respondent, an egocentric network was created. This ego network was used to create a list of the top ten participants based on the highest weighted degree for every survey respondent. In the conversation-based weighted degree-based list, the top ten participants were selected to adjust for noise in the data. In this case, the noise was linked to conversations about general announcements about upcoming events, logistics issues, and queries to the organizers, which may not have anything to do with the strengthening or building ties. Two egocentric networks were generated using Facebook and Twitter data, and the two top ten name lists based on Facebook and Twitter data were compared to the responses.

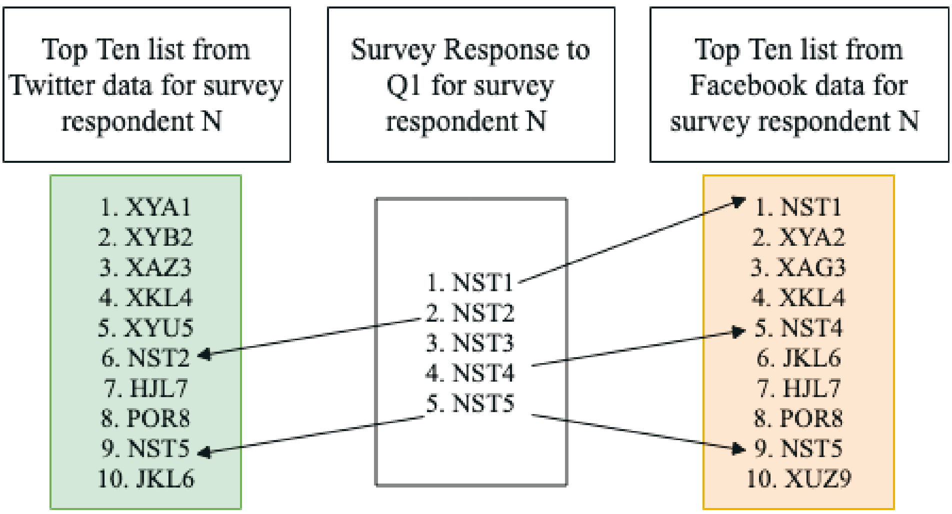

Figure 4.6 Logic for Calculation of individual percentage match.

This comparison was made using the logic shown in Fig. 4.6. For instance, if respondent N responded to question 1 with five participants' names, these names were matched with the names from the top ten list from Twitter and Facebook. The number of names matching each individual was recorded from the top ten lists of

Table 4.3 Correlation between strong ties based on survey and social media data

	Q1	Q2	Q3	Q4
Total number of names received from 24 survey respondents	94	79	77	52
Total number of names identified using the Twitter top 10 list based on the weighted degree of each of the 24 survey respondent	29	26	28	15
Total number of names identified using the Facebook top 10 list based on weighted degree of each of the 24 survey respondent	20	20	16	12
Accuracy in terms of percentage of names identified from Twitter	30.9%	32.9%	36.4%	28.8%
Accuracy in terms of percentage of names identified from Facebook	21.3%	25.3%	20.8%	23.1%
Total number of names received from 24 survey respondents not found in Twitter data	8	6	6	6
Total number of names received from 24 survey respondents not found in Facebook data	30	27	29	17

both Twitter and Facebook of the individual respondent. The process was followed for all responses from each respondent. Table 4.3 presents this aggregated result. The Q1 to Q4 in Table 4.3 refers to the first to fourth questions in the survey (Table 3.7). The first row of Table 4.3 shows the total number of names received from the 24 respondents for each of the questions asked in the survey. These names indicate whom respondents consider their strong ties. The second and third rows indicate the total number of names identified from Twitter and Facebook data. The fourth and fifth columns provide the total number of names from respondents not found on Twitter and Facebook for each survey question.

4.3 Experiment: Kickstarter

4.3.1 Experiment: Role of implicit social ties in crowdfunding project success

Logistic regression was used for data analysis using the R language. The Kickstarter project category was used as a control variable in the data analysis. Table 4.4 provides descriptive statistics of the different measures which were used for testing the hypothesis. This table provides the maximum, minimum, and average values of the different measures in the data.

Table 4.5 provides the results of the study. It can be seen from the table that the

Table 4.4 Descriptive Statistics

Measure	Maximum	Minimum	Mean
Number of Tweets	5049	1	80
Number of Retweets	605	0	19
Number of Mentions	4142	0	96
Avg. Retweet count per potential backer	935	0	3.5
Avg. Follower count per potential backer	1618405	0	12360

structural dimension, which is measured using the degree centrality and betweenness centrality, is found to be significantly related to the crowdfunding project success. Thus, hypothesis X4 is supported. However, even though the measure Average Number of Followers Per Potential Backer used to measure the reciprocal services dimension is significantly related to project success, the relationship is negative. So hypothesis X2 is not supported. Also, since the other measures used to measure the emotional intensity dimension, the intimacy dimension is not significantly related to the crowdfunding project success. Thus, hypotheses X1 and X3 are not supported.

4.3.2 Experiment: Role of implicit social ties of project owners and project backers in degree of success of crowdfunding project

Table 4.6 provides descriptive statistics for some measures used in this study's hypothesis testing. This table depicts the maximum, minimum, and average values of the different measures in the data. To test our research instrument's multilinearity, the maximum variance inflation factors (VIFs) were calculated. All VIFs were below the cutoff value of 5 for regression models, indicating that multilinearity is not a concern in this study (James et al. 2013).

Linear regression was used for data analysis with the statsmodel Python package. The sum of the total number of tweets, retweets, and mentions for each crowdfunding project and project category were set as control variables. As Model 1 shows, project owners' degree centrality is positively associated with crowdfunding project success, and project owners' betweenness centrality has a negative impact when test-

Table 4.5 Results of logistic regression

Dimension	Predictors	coefficient	t-value	*p-value*	Hypotheses test
Emotional Intensity	# of Tweets	4.538e-04	$0.491^{n.s.}$	0.623	X1 rejected
Intimacy	# of Mentions	$2.458e^{-04}$	$0.267^{n.s.}$	0.789	X2 rejected
	Average # of Followers PerPotential Backer	$-2.672e^{-06}$	-1.979 *	0.047	
Reciprocal Services	# of Retweets	$-2.561e^{-03}$	$-0.974^{n.s.}$	0.329	X3 rejected
	Average # of Retweet Per Potential Backer	$-3.148e^{-04}$	$-0.136^{n.s.}$	0.892	
Structrual	Betweenness Centrality of project owner	$3.573e^{+00}$	5.501^{***}	$3.79e^{-08}$	X4 supported
	Degree Centrality of project owner	$2.103e^{-01}$	2.229^{*}	0.025	
Project categorization		Control variable			
McFadden's Pseudo R^2		0.135			

Note:
1. Dependent Variable: Success of crowdfunding project
2. ***: $p < 0.001$; **: $p < 0.01$; *: $p < 0.05$; n.s.: Not significant.

Table 4.6 Descriptive Statistics

Measure	Minimum	Maximum	Mean
Retweet Count Of Project Owner	0	66.250	0.990
# Of Follower Project Owner	0	960041.0	2613.978
Retweet Count Of Potential Project Backer	0	1403.0	4.195
# Of Follower Potential Project Backer	0	$2.786937e^{+06}$	$1.554865e^{+04}$
Betweenness Centrality Of Project Owner	0	1.0	0.065
Degree Centrality Of Project Owner	0	4.0	0.661
Betweenness Centrality Of Potential Project Backer	0	0.250	0.012
Degree Centrality Of Potential Project Backer	0	4.0	0.394
Degree of Success of Crowdfunding Project	0	2.049	438.140

ing only the impact of the four social-tie variables for project owners on crowdfunding projects' success. When testing only the four social-tie variables' impact for project backers on crowdfunding projects' success, project backers' degree centrality was found to positively influence crowdfunding projects' success, whereas retweet counts, numbers of followers, and project backers' betweenness centrality has no significant impact (see Model 2). In Model 3, we tested the full model and found that project owners' degree centrality is positively associated with crowdfunding projects' success, supporting Y3a. Moreover, project owners' betweenness centrality has a negative impact, and the other variables have no significant impact, thus not supporting Y1, Y2, Y3b, or Y4. Table 4.7 provides the results for Model 1, Model 2, and Model 3.

A post hoc analysis was conducted to examine whether a crowdfunding project's funding goal can moderate social ties' impact on crowdfunding project success. We separated projects into two datasets based on the mean value of all project funding goals. The projects whose funding goals exceeded this mean value were referred to as big projects, whereas the mean value exceeded small projects' funding goals, as Model 4 and Model 5 show. The retweet counts, numbers of followers, degree centrality, and betweenness centrality of both project owners and project backers have no significant impacts on crowdfunding projects' success for big projects. For small projects, project owners' degree centrality is positively associated with crowdfunding project success, project owners' betweenness centrality has a negative impact, and the other variables have no significant impact. These results are shown in Table 4.8.

Table 4.7 Linear regression for all studied Kickstarter projects' degrees of project success

	Model 1	**Model 2**	**Model 3**
Intercept	2.657^{**}	3.013^{**}	$2.800^{n.s.}$
Retweet Count Of Project Owner	$0.129^{n.s.}$		$0.046^{n.s.}$
# Of FollowerProject Owner	$0.559^{n.s.}$		$0.558^{n}.s.$
Degree Centrality Of Project Owner	3.758^{***}		3.504^{***}
Betweenness Centrality Of Project Owner	-2.028^{*}		-2.642^{*}
Retweet Count Of Potential ProjectBacker		$-0.168^{n.s.}$	$-0.029^{n.s.}$
# Of Follower Potential Project Backer		$-0.473^{n.s.}$	$-0.355^{n.s.}$
Degree Centrality Of PotentialProject Backer		1.994^{*}	$-1.448^{n.s.}$
Betweenness Centrality Of Potential Project Backer		$-0.503^{n.s.}$	$-0.306^{n.s.}$
R^2	0.016	0.012	0.018
Adjusted R^2	0.008	0.003	0.007
Number of Observations	2161	2161	2161

Note:
Dependent variable: degree of success of a crowdfunding project.
***: p <0.001; **: p <0.01; *: p <0.05; and n.s.: not significant.

Table 4.8 Linear regression test results for big and small projects' degrees of project success

	Model 4 (Big Projects)	Model 5 (Small Projects)
Intercept	$0.066^{n.s.}$	2.672**
Retweet Count Of Project Owner	$-0.458^{n.s.}$	$0.330^{n.s.}$
# Of FollowerProject Owner	$1.022^{n.s.}$	$0.547^{n.s.}$
Degree Centrality Of Project Owner	$0.543^{n.s.}$	3.701***
Betweenness Centrality Of Project Owner	$-0.224^{n.s.}$	-2.594*
Retweet Count Of Potential ProjectBacker	$-0.643^{n.s.}$	$-0.005^{n.s.}$
# Of Follower Potential Project Backer	$-0.028^{n.s.}$	$-0.298^{n.s.}$
Degree Centrality Of Potential Project Backer	$-1.042^{n.s.}$	$-1.584^{n.s.}$
Betweenness Centrality Of Potential Project Backer	$1.572^{n.s.}$	$-0.386^{n.s.}$
R^2	0.017	0.012
Adjusted R^2	0.008	0.003
Number of Observations	148	2013

Note:
Dependent variable: degree of success of a crowdfunding project.
***: $p < 0.001$; **: $p < 0.01$; *: $p < 0.05$; and n.s.: not significant.

5 DISCUSSION

The following chapter discusses the meaning of the empirical findings towards identifying and role of implicit social ties in different contexts. Accordingly, this chapter discusses how the three different studies help address the three research questions of this dissertation and how these studies' findings are related to the earlier studies.

5.1 Case: Indie Game Developer Community

The following was the first sub-question (RQ1): *How can different kinds of implicit social ties be identified using the publicly available social data and big social data ?* This study showed that the different kinds of implicit social ties could be identified using an edge centred approach. This study aimed to find relevant actors (nodes) and their reciprocal interactions (edges) from publicly available large social media datasets in the context of the tie strengthening process. This study sought to overcome the limitations of earlier studies that used techniques that were not appropriate (due to legality or privacy issues) or studies that used measurements that may not accurately reflect individuals' true social connections. In this study, the edge-centered approach was demonstrated with data from Facebook and overcame many of these earlier limitations.

The edge-centered approach in tie strength research initially allows researchers to build a network of reciprocal interaction from open social media data without compromising current legal or privacy concerns. In addition, the big data paradox and noise removal fallacy are mitigated by getting rid of "low frequency" connections between individuals. Strong ties take time and frequent interaction to form, so low-frequency connections are less relevant from a strengthening perspective (Granovetter 1973; Walther 1992).

An edge-centered approach offers many interesting possibilities, for example, integrating large-scale SNA with text mining and machine learning algorithms that can automatically analyze and discover patterns in a large amount of textual data (Debortoli et al. 2016). Any social media dataset containing discussions could be analyzed with an edge-centered approach. Analysis of social network data and message content has been used, for example, to identify online leadership (i.e., the language used and the sociability of individuals) (Faraj, Kudaravalli and Wasko 2015). However, studies analyzing the content of messages often use manual coding (such as the study mentioned above), not automated text mining or machine learning algorithms (Debortoli et al. 2016).

This approach allows researchers to test well-established sociological theories and assumptions from human behavior. For instance, Granovetter 1973 mentioned that "the strength of a tie is (probably a linear) combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie". There is also a plethora of similar assumptions in other theories that could be tested by analyzing many different communication patterns in social networks.

As an example, reciprocity is a critical "starting mechanism" for establishing and strengthening social relationships in the early stages. Simple social exchanges, such as expressing respect and appreciation (Gouldner 1960), can also be identified from text-based interactions. In addition to enabling significant theoretical contributions, this kind of confirmatory research (theory testing) could also guide organizations to manage employee interactions best. In the future, it would be interesting to explore what kinds of communication patterns endure (and create stronger bonds among individuals) and what types fade away.

Moreover, this approach complements research methods that rely on electronic data (likes, links shared, etc.) or self-reported data (questionnaires, etc.). These methods usually suffer from inaccuracies, recall- and cognitive biases, or perception errors (Wuchty 2009).

5.2 Case: CMAD

The following was the second sub-question (RQ2): *How can the potentially useful implicit social ties be identified using social data and big social data in different knowledge work context like professional conferences, research conferences and other similar contexts?* As a response to this research question, this study utilized the publicly data related to an event for the identification of implicit social ties. In this study, different kinds of analysis methods were used to identify social ties from publicly available data about an event. The social media conversation data of Twitter and Facebook about the event was used to create two different people networks. The combination of cluster analysis of the Twitter network and content analysis of the clusters led to the identification of theme-based sub-communities within the CMAD 2016 event based on the Twitter network. In this way, the different structural holes could be identified based on the network structure. Figure 4.4 shows an example of the non-redundant bridging nodes between the different subcommunities by illustrating nodes like W, N1, D2, and A3, among others. Those nodes would be regarded as actors bridging structural holes and deemed to comply with Burt's definition.

Individual survey responses for question 6 revealed that survey respondents' most novel sources of information were the non-redundant bridging nodes based on the network in Figure 4.4. The identified actors in Figure 4.4 and Table 4.2 were bridging the structural holes pointing towards weak ties or potential weak ties. These weak ties serve as the bridging ties linking the various possible sub-communities of the entire CMAD community identified in Table 4.1.

Some of the analyses done in this study involved using survey data. However, the survey was answered by a small number of respondents, making it difficult to draw statistically significant results. Hence, this also provides the following propositions that strive to enhance the current understanding of the research objective of this study.

Proposition 1: Identifying the social ties from social media data in an event accurately depends on the purpose and pattern of use of the particular social media platform.

According to the temporal distribution of social media data collected (Fig. 4.2 & 4.2), Twitter was used most during the CMAD2016 event day, while Facebook was more popular before the event start. In addition, from the force-driven network of people (Fig. 4.5 & 4.4), one can observe that Facebook has an apparent central node. In contrast, Twitter does not have a central node. A likely data-driven explanation for these findings would be that Facebook Page might have been used for planning the event. In contrast, Twitter was used only during the event for maintaining ties or building new ones. An analysis of the tweets' descriptive analysis supports this preliminary explanation. It's a novel approach academically since when you rely on publicly available social media data, it's essential to understand the purpose and pattern of how social media platforms are used. This is because the only source for deriving the implicit relationships is the content of the textual data. In an event, different social media channels may be used for very distinct purposes. In the absence of consideration of this aspect when deciding which social media channel to use for evaluating tie strength, an irrelevant estimate of tie strength would result from selecting the wrong one. This would lead to incorrect or irrelevant identification of the social ties.

Proposition 2: The structure of the implicit social network can reveal possible weak ties, provided the selected social media channel is used to maintain and build ties.

According to the responses to survey question 6 (Table 3.7) and the force-driven network of people based on tweets (Fig. 4.4), 80% of respondents do not belong to the same cluster as the people who are their novel sources of information. These novel sources of information were also associated with many event participants from different clusters (Fig. 4.4). Based on the literature, it is known that weak ties are a source of novel information and act as a bridge between diverse people (Granovetter 1973). Therefore, the above empirical findings, concomitantly with existing literature, provide support to proposition 2 statement. The current proposition is academically novel since previous studies have only used explicit relationship social media data to create a network of relationships, which could then be used as a basis for identifying ties (e.g. Backstrom and Kleinberg 2013). However, in this study, only implicit relationships derived from the event participants' communication over social media were used. In the case of an event, such data is easily accessible, while explicit relationship information is nearly impossible to find. If this proposition is validated in the future, it could represent a new method for identifying weak ties, which would be highly relevant for building collaboration tools, such as social recommendation systems based on tie strength.

Proposition 3: The weighted degree from implicit relations from social media data can correlate with tie strength, especially strong ties, provided the chosen social media channel is used to maintain and build relationships.

This study found preliminary evidence to support this proposition. From Table 4.3, it is possible to identify survey participants' perceived strong ties with an accuracy of about 30% in Twitter data and about 20% in Facebook data. This accuracy in predicting the perceived strong ties is good because identifying the strong ties was executed only using social media content. No other explicit relationship data from social media was used to either identify the existence of the tie or for the specific identification of strong ties. This study used only the weighted degree calculated from implicit networks derived from social media content. Table 4.3 provides preliminary empirical evidence for this proposition based on the results of this analysis. This proposition is novel from an academic perspective since previous studies on the topic (e.g., Ahn and Park 2015; Aral and Walker 2014; Gilbert and Karahalios 2009) have used measures that consider explicit online relationships or collect data from private social media accounts. Nonetheless, in this study, the measure used for tie strength evaluation was based on publicly available social media data. Practically, this aspect is critically important. In most scenarios, it is almost impossible to obtain explicit relationship data or private user data about participants from social media. In contrast, textual content (e.g., Tweets in the case of Twitter, Text from open Facebook pages in the case of Facebook) related to the event is relatively easy to collect. Therefore, measures, which can evaluate tie strength from this type of social media data, will be helpful while developing tie strength-based conference recommender systems.

5.3 Experiment: Kickstarter

The following was the third sub-question (RQ3) of this study: *How can the potentially useful implicit social ties, identified from existing social data and big social data, be used in different professional contexts like business decision making or business phenomena?* To answer this research question, implicit social ties were identified from social data related to crowdfunding projects and also the role of these identified implicit social ties towards crowdfunding project success was studied.

5.3.1 Experiment: Role of implicit social ties in crowdfunding project success

This part of the study explores the impact of implicit social ties from social media on the success of crowdfunding projects. First, this study's empirical results suggest a significant link between the structural dimension of tie strength and the success of crowdfunding projects. By measuring both the betweenness centrality and the degree centrality of the project owner, the structural dimension of tie strength was determined. Study findings indicated that project owners' centrality and degree centrality in social media were directly related to crowdfunding project success. A more substantial impact was found on the success of a crowdfunding campaign by considering the betweenness centrality. By measuring betweenness centrality, one measures the extent to which a user falls on the shortest path between two other users in the network, i.e., the extent to which the user plays the role of bridging between pairs (Hansen et al. 2020). In this context, a bridge is a (strong or weak) relationship where there is no indirect connection through third parties or other users Burt 2004. Users with high betweenness centrality form non-redundant, often weak ties among less connected users. In the context of Twitter, these users facilitate access to other users in different networks that would otherwise not be possible (Hansen et al. 2020). Additionally, the degree centrality measure was found to be statistically significant for crowdfunding project success. A user's degree of centrality is determined by the number of unique connections in a network. In Twitter, degree centrality measures user engagement with other users and their contents. It indicates actual attention to content and actions users took to share information. Twitter users with a high degree centrality act as conversational hubs (Hansen et al. 2020).

According to the results of this study, the crowdfunding project owner's social network structure could drive potential backers to crowdfunding and help spread crowdfunding project information to their social networks, resulting in the crowdfunding project's success. This could be since the implicit ties between crowdfunding project owners in social media provide potential backers with information about the project initiators. Potential project backers use this information to determine whether to fund the crowdfunding project based on their economic interests. Furthermore, it facilitates disseminating information about the crowdfunding project to a broad audience through the owner's weak social ties.

Lastly, the empirical results of this study indicate that the number of tweets, the number of mentions, the average number of followers per potential backer, the number of retweets, and the average number of retweet per potential backer for a crowdfunding project on Twitter are not positively associated with the success of a crowdfunding project. Based on the findings, the emotional intensity, intimacy, and reciprocal service dimensions of tie strength in social media for a crowdfunding project are not directly associated with its success.

5.3.2 Experiment: Role of implicit social ties of project owners and project backers in degree of success of crowdfunding project

This study investigates the effects of project owners' and potential project backers' social ties on crowdfunding projects' degree of success. Study results show that project owners' degree of centrality is significantly related to crowdfunding projects' success. A node's degree centrality is simply a measure of the number of connections that connect it to other nodes in the network. This measure helps to identify highly connected individuals or who can quickly connect with a wider network (Borgatti, Everett and Johnson 2013). In Twitter's context, degree centrality captures users' engagement with other users and their content. Users with a high degree centrality serve as conversational hubs (Hansen et al. 2020). By having a high degree centrality, a crowdfunding project owner could distribute the project information to a broader network of project backers, potentially increasing project funding.

The relationship between the project owners' betweenness centrality and the degree of a crowdfunding project's success is statistically significant, but this relationship is negative.The Betweenness centrality of a node measures how often it falls on the shortest path between two other nodes(Borgatti, Everett and Johnson 2013; Freeman 1978). In general, betweenness indicates the possibility of controlling flows through a network–that is, acting as a gatekeeper or toll collector. In general, nodes with high betweenness can color or distort information as it is passed along. In addition, the ability to exploit a high betweenness position varies inversely with the ease with which nodes can form new social ties (Borgatti, Everett and Johnson 2013). In the Twitter context, a user's high betweenness centrality could indicate that they have access to users from other clusters or simply that they are at the edges of both clusters (Hansen et al. 2020). Due to their high betweenness centrality, project owners may be able to control information flows but, at the same time, may have difficulty building new social ties with other individuals. It would reduce their information flow to more disparate networks, decreasing the chances of attracting more project backers and receiving greater funding.

According to this study, the number of retweets and followers of project owners has no impact on crowdfunding projects' success. Retweeting involves redistributing content across a user's social network. In contrast, followers are used to estimating the size of an individual's social network. While project owners may have more retweets or followers, this might not reflect higher interaction or engagement with their social circles in crowdfunding projects. Therefore, being unable to attract additional project backers and funding for the project.

In this study, retweet counts, number of followers, degree centrality, and betweenness centrality of potential project backers were found to have no statistically significant impact on crowdfunding project success. The possible reason for this result is that, although backers are part of a crowdfunding project's social network, they do not directly provide any detailed information about the project. The project owner is still the only source of information for potential backers of crowdfunding projects. Furthermore, potential backers cannot directly interact with or engage with project owners. They cannot control the flow of information related to a crowdfunding project. Consequently, they have a limited ability to attract new funding for a crowdfunding project.

According to the post hoc analysis of this study, the study's model can explain small projects but not large ones. Initial findings suggest that social ties of project owners and potential backers have no bearing on big crowdfunding projects' degrees of success (see Table 4.7). To reach their funding goals, big projects require extensive funding, and these types of projects must attract large amounts of funding to succeed. The social ties of project owners may provide a relatively small amount of funding compared to a project's overall funding goal. In order to reach their crowdfunding funding goals, such projects need to appeal to as many potential backers as possible who aren't already familiar with the project owners or the existing project backers. As a result, the social ties of project owners and potential backers play a minimal role in helping to achieve crowdfunding project goals. In contrast, for small projects, social ties play a role in crowdfunding success since their funding goal targets are relatively small and do not require many potential contributors. Hence, small projects can utilize project owners' social ties to gain funding. The findings of this study are in line with some earlier studies (Cumming, Leboeuf and Schwienbacher 2020; Koch and Cheng 2016; Mollick 2014) which found that, for large funding goals, social ties did not influence crowdfunding success.

The results of the study have important practical implications. According to the study, degree centrality is positively correlated with the success of crowdfunding projects. This finding suggests that project owners should engage with other social media users to help promote their crowdfunding projects. Moreover, they should directly interact with other users on social media as much as possible. Direct interaction with other social media users will likely increase project funding.

Second, the results indicate that project owners' betweenness centrality negatively correlates with their projects' success. In practice, the negative relationship implies that although project owners should engage with as many other social media users as possible, they should also evaluate whether they are the only point of contact for different groups of people. A project owner must consciously engage with other social media users (potential backers) from disparate networks (for example, networks based on different locations or interests) instead of remaining the primary user connecting disparate networks. In addition, owners of projects should consider creating a platform that connects social media users from disparate networks, thus providing a chance for multiple points of contact among different social media users.

Lastly, the results show that project funding goal sizes also influence project owners' and potential backers' social ties in terms of crowdfunding success. Social ties are only helpful in crowdfunding projects with small funding goals, according to this study. Social ties of project owners and potential project backers do not influence crowdfunding success for large projects. This finding means that project owners should maximize the use of their social ties in crowdfunding projects with small funding goals since these ties can be crucial to achieving funding goals. In contrast, project owners of large crowdfunding projects should devise strategies to reach as many unknown potential backers as possible with information about their project. Projects with such large funding goals are more likely to meet their funding goal if more unknown potential backers are attracted to the crowdfunding project.

5.4 Synthesis of different case studies and experiment

The previous three sections help in understanding how the three different sub-studies address the different research questions of this dissertation. This section provides an overall perspective about how these three sub-studies are interlinked and complement each other to achieve the overall objective of this dissertation. This is also illustrated in the Fig. 5.1

The focus of the first sub-study Case: Indie Game Developer was towards developing an approach which could be used for identification of implicit social ties from publicly available social data and big social data. This study focused on developing

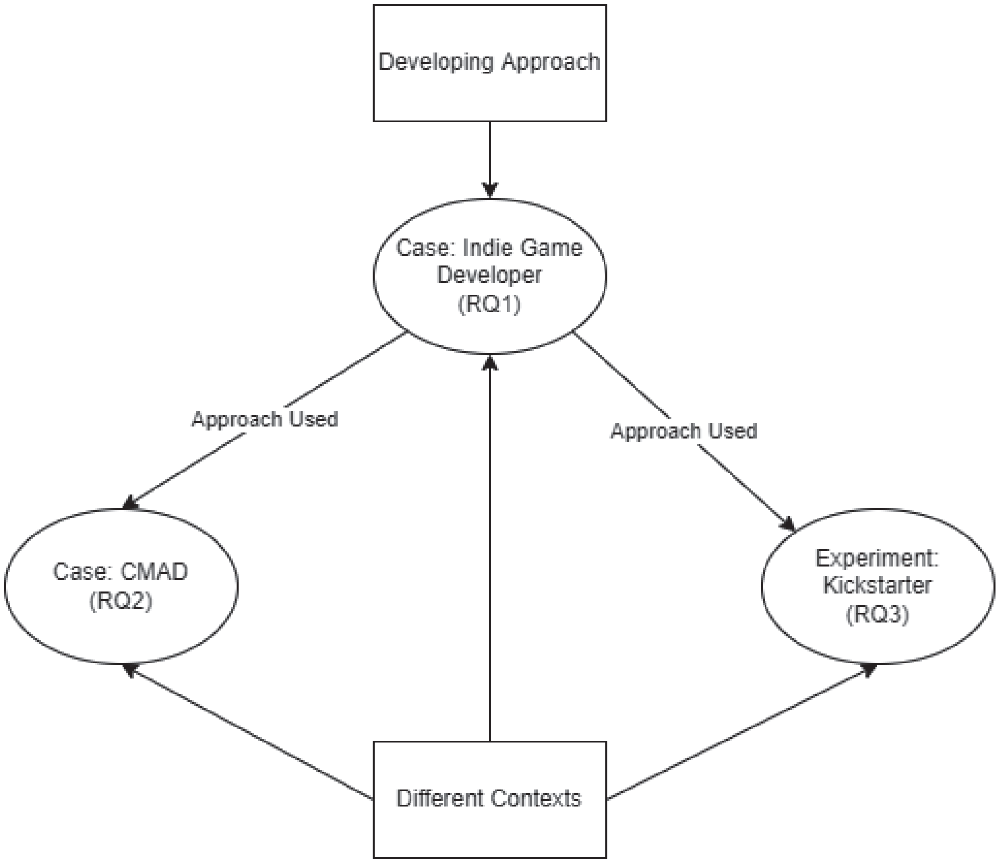

Figure 5.1 Role of the three sub-studies in the dissertation

an approach which could be used for identification of implicit social ties from publicly available social data and big social data. The study resulted in the development of an approach for the identification of implicit social ties from publicly available social data and big social data. The developed approach could be adapted to identify the implicit social ties in different contexts. In the context of this dissertation the sub-study Case: Indie Game Developer provided the first step towards the achieving the overall objective of the dissertation. At the same time this sub-study also provided one of the contexts where the identification of implicit social ties could be done. This is also shown in Fig. 5.1.

The subsequent two sub-studies Case: CMAD and Experiment: Kickstarter were able to utilise the approach developed in sub-study Case: Indie Game Developer for the identification of implicit social ties from publicly available social data and big social data. Both the sub-studies Case:CMAD and Experiment:Kickstarter enabled studying the role of implicit social ties in two very different contexts. This is also illustrated in Fig. 5.1.

From the perspective of this dissertation the different sub-studies played different roles towards achieving the overall objective of this dissertation. The sub-study Case: Indie Game Developer enabled the development of an approach to identify implicit social ties from publicly social data and big social data. The sub-study Case: CMAD utilised the approach developed in Case: Indie Game Developer for the identification of implicit social ties from publicly available social data in the context of knowledge work like professional conferences. The sub-study Case: CMAD focused in-depth on the identification of useful implicit social ties in the context of knowledge work context like professional conferences by using the single case study strategy leading to more comprehensive understanding. The sub-study Experiment: Kickstarter also utilised the approach developed in Case: Indie Game Developer for the identification of implicit social ties from publicly available social data in the context of crowdfunding success. This sub-study focused on understanding the role of implicit social ties towards crowdfunding success. The sub-study used the experiment strategy to understand the relationship between the dependent and independent variable and obtain statistically significant results. Hence, at a broader level the three different sub-studies provided three different contexts for the identification of implicit social ties. At the same time the approach developed in one of the sub-study was adapted and used in the other two sub-studies in order to further understand the role of implicit social ties in different contexts. Thus, the three sub-studies together complemented and provided synergies for achieving the overall objective of this dissertation.

6 CONCLUSIONS

6.1 Conclusions and academic contributions to the research questions

The conclusions and contribution of this dissertation project to answering the research questions is summarized in this section. All the studies in the dissertation project contributed to the intersection of social data, social ties, different theoretical approaches and different contexts as shown in Figure 1.1.

Table 6.1 presents the major contributions of the dissertation in correspondence with the research questions of the dissertation. The most important area of academic contributions of each sub-study of this dissertation are also highlighted in the Table 6.1. More detailed contributions are presented in the following parts of this subsection.

Since the inception of social media, a lot of studies have been done to calculate tie strength to identify different kinds of social ties using social media data (e.g. Fogués et al. 2013; Gilbert and Karahalios 2009; Kahanda and Neville 2009). However, most of these studies have used many different methods of data collection from social media which can no longer be used due to changes in data access policies of social media platforms. This was also highlighted in section 2.2.2 and 2.3.1. At the same time in the recent years, the collective process of production, consumption,and diffusion of information on social media are starting to reveal a significant portion of human social life (Davis et al. 2016). This has led to two major problems related to social media data: (1) big data fallacy and (2) noise removal fallacy. Big data fallacy refers to the condition where there is access to large amount of social media data which is mostly publicly available, yet we may have very little detail about an individual from this big data directly. On the other hand, noise removal fallacy refers to firstly the removal of noisy data from social media data may actually remove valuable infor-

Table 6.1 Major conclusions and academic contributions of the studies towards the research questions of the dissertation

Research Question	**Main academic contributions of the studies towards the research questions** (for more detail related to contribution to earlier research, see text in section 6.1)
RQ1: How can different kinds of implicit social ties be identified using publicly available social data and big social data?	Case: Indie Game Developer Community **contributes to the research related to using publicly available social media data for identifying different kinds of social ties**. This study shows that the approach of using reciprocal interaction-based social networks (type of implicit social network) derived from social media data can be used to filter out the relevant implicit social ties and be used to identify different kinds of implicit social ties.
RQ2: How can the potentially useful implicit social ties be identified using social data and big social data in different knowledge work context like professional conferences, research conferences, and other similar contexts?	Case: CMAD **contributes to the research related to the identification of different kinds of implicit social ties from social media data in the context of events**. This study provides empirical evidence that Structural Hole Theory can be used to identify the useful weak ties from implicit social networks in the context of an event. The study provides preliminary evidence that the structure of the implicit social networks can be used for the identification of weak ties while the network measure - weighted degree calculated from implicit social network correlates with tie strength especially in the case of strong ties in the context of events.
RQ3: How can the potentially useful implicit social ties, identified from existing social data and big social data, be used in different professional contexts like business decision-making or business phenomena?	Experiment: Kickstarter **contributes to the research related to the role of social ties towards crowdfunding success**. The first part of the study provides empirical evidence that the implicit social ties (derived from social media data) positively affect the crowdfunding project's success. The second part of the study contributes to research about the role of social ties of project owners and project backers towards the degree of success of crowdfunding projects. This study tests some hypothesis which uses network measures derived from implicit social networks (based on interaction data about crowdfunding projects on Twitter) involving project owners and potential project backers. The study provides empirical evidence that the social ties of project owners impact the degree of crowdfunding success.

mation and secondly, the definition of noise is complicated and depends completely on the task at hand. (Zafarani, Abbasi and H. Liu 2014) However, there is very little research on how to identify relevant social ties from publicly available social data while also considering the above mentioned social media data related issues.

Case: Indie Game Developer Community Case: Indie Game Developer Community furthers the understanding about identifying different kinds of social ties from publicly available social media data by creating reciprocal interaction network based on data about an Indie Game Developer Group on Facebook. The results of this study show an edge-centred approach by creating an interaction network which is composed of nodes (representing the actors) and edges (representing the interaction between nodes). This edge-centred approach enables filtering out and identifying the relevant actors from large social media datasets based on the interaction. At the same time this approach enables the tie strength researchers to zoom in or zoom out from the large social media dataset based on the current tie strength related task at hand.

The approach developed in this study complements the existing research methods that rely on simple digital measures (e.g.likes, links shared) or self-reported data (e.g. questionnaires) that may sometimes suffer from inaccuracy, recall- and cognitive biases, or errors of perception (Wuchty 2009). This study at a broader level helps in the identification of patterns of communication and contributes to the existing stream of literature (e.g. Lazer et al. 2009; Pentland 2015).

Case: CMAD Social media has provided a new way of networking with other people not only in private life but even in co-located professional events like professional conferences (A. Zhang, Bhardwaj and Karger 2016). One of main the aims of the participants in such professional events is to meet new people who may share similar interests or may provide relevant information and are potentially useful for them (Xia et al. 2013). However, there is a very limited time in such events for such opportunities.(Oppermann and Chon 1997; Severt et al. 2007; A. Zhang, Bhardwaj and Karger 2016) This lack of time acts as potential challenge and limits the chance of meeting potentially useful contacts. In order to avoid this challenge, some studies have shown that participants use social media as a channel for networking (Ross et al. 2011). Use of social media for networking and information seeking has been shown to be one of the most important reasons for social media use in a professional event setting like professional conference (Ebner et al. 2010; Reinhardt, Ebner and Beham 2009; Ross et al. 2011). This has led to the emergence of a stream of literature which is focused specifically on events. (Ahmed et al. 2014; S. Liu, B. Wang and Xu 2017; Tong, She and Meng 2016) Many of the studies in this area have focused on recommending different events to the members/ users of the event based social network sites (e.g. X. Li et al. 2017; She et al. 2016; Tong, She and Meng 2016; Yu et al. 2015). A study has also been done with objective of recommending event participants to other event participants but these recommendations were restricted to the precondition that participants were friends (Y. Lu et al. 2016). Hence, even this study relied on using explicit relationship data which is very difficult or no longer possible to get from most social media platforms as highlighted in Publication 4. In the past, there has been a study to evaluate tie strength in the context of an event, but this study did not use any real empirical data like social media data from a conference (Gupte and Eliassi-Rad 2012). In order to enhance networking and automate the process of participant recommendations in an event, intuitively it would make sense to incorporate the concept of social media based social ties while developing such recommendation systems. Thus, there is a need to understand how to identify social ties using publicly available data in the contexts of professional events. However, there is very limited research on the identification of different kinds of social ties using the publicly available social media data in the context of an event.

Case: CMAD increased the understanding about the identification of social ties from social media data in the context of an event (RQ2) by utilizing the social media data from two different platforms - Facebook and Twitter about the professional event CMAD2016 for identifying different kinds of social ties. Case: CMAD furthers the understanding about identifying social ties from social media data in the context of an event (RQ2) by utilizing the theory of structural holes to identify weak social ties and also by using the concept of tie strength to identify different kinds of social ties. This study contributes to the understanding about the use of social media related data sources for identifying actual or potential weak ties in an event context which may be used e.g. to automate the process of identifying different ties (e.g. Kahanda and Neville 2009; Xiang, Neville and Rogati 2010) to connect professionals in events. Publication 1 also adds to the present literature about the different interpersonal measures used for identification of social ties (e.g. Fogués et al. 2013; Gilbert and Karahalios 2009; M. Granovetter 1983; Marsden and Campbell 2012) by using network measures based on the concept of structural hole theory.

Case: CMAD shows that the purpose and pattern of use of a particular social media channel in an event impacts how accurately the social ties can be identified from that social media data. The study provides preliminary evidence that the structure of the implicit social network can reveal the possible weak social ties. This result is novel and academically contributes to the existing literature related to the creation and use of social media based relationship network for identifying different social ties. However, unlike previous studies which utilized explicit relationship data from social media (e.g. Backstrom and Kleinberg 2013), this study only uses the implicit social network. The study also provides preliminary evidence that the weighted degree based on the implicit social network can correlate with tie strength, especially in case of strong ties. Thus, this study at a broader level also adds to the existing literature about the different measures used to calculate tie strength from social media data. However, unlike previous studies which created these measures using private or explicit relationship data s (e.g. Ahn and Park 2015; Aral and Walker 2014; Gilbert and Karahalios 2009) , the measures developed in this study used only publicly available social media data and did not use any explicit relationship data.

Experiment: Kickstarter Over the past decade, crowdfunding has become a novel and essential method of raising funding to carry out projects which have previously not been possible easily. According to Mollick 2014, "crowdfunding is a novel method for funding a variety of new ventures, allowing individual founders of for-profit, cultural, or social projects to request funding from many individuals, often in return for future products or equity." In other words, crowdfunding provides an alternate way of connecting people who need resources to carry out a project or an idea with those who are willing to anchor them to start a project or a business through digital platforms (Madrazo-Lemarroy, Barajas-Portas and Labastida Tovar 2019). The different digital crowdfunding platforms have given rise to different crowdfunding models. Broadly there are four kinds of crowdfunding models, including donation based, investment based, crowdsourcing and lending groups based and reward based. Over the years, extant literature has researched the different factors related to crowdfunding projects' success. Some of these studies (e.g. J. S. Hui, Gerber and Greenberg 2012) have focused on the crowdfunding process where as another research stream (e.g. Gerber and J. Hui 2013) has focused on investigating the motives and deterrents for participation in crowdfunding platforms. Another stream of literature has focused on investigating the effect of social capital of fundraisers (e.g. Kang, Jiang and C. H. Tan 2017). Furthermore, research has mainly examined design strategies for crowdfunding projects, including design strategy dealing with how to design the project content (see Mollick 2014), reward schema (see N. Zhang, Datta and Kannan 2014), and social media usage related to the crowdfunding project. Current research has shown that social media has become an important factor contributing to the success of crowdfunding project. Most of these studies have used social media based measures like number of Facebook friends, Twitter followers, which rely on explicit relationship data from social media (Hong, Hu and Burtch 2018; Mollick 2014). A few studies (e.g. Borst, Moser and Ferguson 2018; Kang, Jiang and C. H. Tan 2017) have looked into the impacts of different dimensions of social ties on the crowdfunding project success. However, there is little or no research on the role of implicit social ties derived from social media data on the success of crowdfunding projects.

Experiment: Kickstarter increase the understanding about the role of implicit social ties derived from social data in the context of a business context or decision making (RQ3) by analysing the role of implicit social ties derived from Twitter data on the success of the crowdfunding projects of the reward based crowdfunding platform Kickstarter.com. The first part of the study furthers the understanding about the role of implicit social ties derived from social media data in a business or decision making context (RQ3) by analysing the role of implicit social ties derived from Twitter data in the success of the crowdfunding projects from the crowdfunding platform Kickstarter.com. More specifically, this study analysed how the four different dimensions of tie strength of implicit social ties derived from Twitter data about the crowdfunding project impact the crowdfunding project success for projects from Kickstarter.com. The results of this study show that the implicit social ties do impact the success of crowdfunding project. The study results show that the structural dimension of tie strength has statistically significant impact on the crowdfunding project success. At a more general level, this study contributes to the stream of literature (e.g. **Clauss2020INCREASINGCROWDFUNDING**; Hong, Hu and Burtch 2018; Kromidha and Robson 2016; Mollick 2014) about the impact of network size and network structure derived from the explicit ties from social media. The second part of the study furthers the understanding about the role of implicit social ties derived from social media data in a business or decision making context (RQ3) by analysing the role of implicit social ties (derived from Twitter data about the crowdfunding projects) of the potential project backers and project owners on the degree of success of crowdfunding projects (projects from Kickstarter.com). More specifically, this study developed four different hypothesis related to the role of social ties of project owners and potential project backers on the degree of success of crowdfunding project. Two of these hypothesis used network measures which were calculated from implicit social networks derived from interaction data about crowdfunding project as a proxy for the social ties of project owners and potential project backers. The results of this study show that the social ties (derived using social media data) of project owner affect the degree of success of the crowdfunding project. This study contributes to the stream of literature related to the role of social ties of project owners and project backers on the crowdfunding success (Agrawal, Catalini and Goldfarb 2015; Borst, Moser and Ferguson 2018; Kang, Jiang and C. H. Tan 2017; Y. H. Tan and Reddy 2021) .

The Table 6.2 shows how the different studies of this dissertation help address some of the different challenges related to using social data to identify social ties, which were discussed in the section 2.3.1 of this dissertation. The Table 6.2 also shows the implication related to the different 'V's of big data, which the different studies of this thesis address. Thus, this dissertation addresses some of the challenges related to the use of social data for the identification of social ties and also provides the implications related to the different aspects of big data addressed by the different studies of this dissertation.

6.2 Managerial implications

The identification of implicit social ties using social data and big social data brings in new challenges which need to be considered. There is a need to understand that many factors like regulations (such as GDPR), platform business model changes and experiments, public exposure of unethical use of social media data can quickly and sometimes even quite unexpectedly impact the data access policy decisions of social media platforms. Along with these factors, there is a long term growing trend in restricting and curtailing the data access by the social media platforms. The researchers and others working in this domain of social tie identification from social media should take these factors into consideration. They should focus more on using methods that are mainly or solely based on available social media discussions and content. One such approach was provided in Case: Indie Game Developer Community. At the same time, there is a need to also focus on look at the existing wide variety and different kind of theoretical approaches and other approaches that have been used in other research areas or with other kinds of social data (e.g. Mattie et al. 2018) and may be suitable for identification of social ties from social data.

The study Case: Indie Game Developer also has many practical and managerial implications. Conversational data from online social networks provides interesting insight into how different "interaction layers" are evolving over time. Information systems (IS) scholars have argued that further research is needed to understand the content of communication at the network level (Faraj, Kudaravalli and Wasko 2015). A visual representation of these interactions significantly enhances sense-making and makes this information more useful for many organizations and decision-makers.

Using large-scale social network analysis in conjunction with ample volumes of

Table 6.2 Challenges related to use of social data for social ties identification and the potential implications towards big data based on the different studies of this dissertation

Study	Challenges Addressed related to identification & use of social ties	Potential implications related to big data (6 'V's)
Case: Indiegame Developer Community	• The study used publicly available social media data about a community. Thus addressing challenge 2 and partially addressing challenge 1 related to using social data. • The study develops an approach based on social network analysis and reciprocal interaction networks to filter out relevant actors. Hence, addressing challenge 4 related to social data.	The study provides an approach that can also be used with a large volume of data. Thus, this study has implications for the Volume aspect related to big data.
Case: CMAD	• The study used publicly available social media data about conversations related to an event. Hence, addressing challenge 2 and partially addressing challenge 1 related to using social data. • The study used network-level measures for the identification of different kinds of social ties. Thus, addressing challenge 4 related to social data.	The study utilizes unstructured text data related to the conversations related to an event from two different social media platforms. Hence, this study has implications for the Variety aspect of big data.
Experiment: Kickstarter	• The study used publicly available social media data related to conversations about crowdfunding projects. Thus, addressing challenge 2 and partially addressing challenge 1 related to using social data. • The study used network-level measures and other social data measures for the identification of different kinds of social ties. Hence, addressing challenge 4 related to social data.	The study utilizes a large volume of social data to identify social ties.This study has implications for the Value aspect of big data as this study utilized a large volume of social data to identify a relatively small number of social ties.

texts leads to insightful findings. As a result, it would be possible to determine what causes excitement in social networking services. For example, who enjoys speaking with whom and what types of communication patterns emerge (emotional content, language formality, and social exchange). This kind of information is helpful in many different contexts to understand the social dynamics of the groups for which the related social media data has been analyzed.

The identification of social ties using social data and big social data in the context of events has to deal with not only the prior mentioned challenges but also some other limitations unique to the context of events. These limitations were discussed in Case: CMAD.

Case: CMAD has many managerial implications which should be considered. First, a social media channel's purpose and patterns of use in an event should be understood before tie strength calculations are conducted. Second, the structure derived from the social media content (i.e. implicit network) could identify the participants who are connecting the clusters of discussion topics and thus could be categorized as potential weak ties. The implications of this would be important for both the organizers of events and the conference recommendation system designers because it would allow them to identify the most diverse and networked attendees.

Third, the kind of event should also be taken into consideration while recommending potentially useful social ties. As an example, in a focused conference with a clearly defined theme and a narrow range of topics covered in it, identifying weak ties may be sufficient to establish a valuable contact as the participants in such an event are very well aware of the topic and theme of the event. The participant and their identified weak ties would likely share the same common interests addressed in such a focused event. On the other hand, there may be a large number of weak ties to be identified in large events with a wide range of themes. However, many of the identified weak ties may not prove useful in this case unless they are prioritized based on certain factors like their interests, skills, and expertise that are of interest to event participants. To address this, further analysis and other sources of data will be required. For example, in the case of large or relatively widely focused academic conferences (like HICSS, CHI, or Academic Mindtrek), the bibliographic data of the event participants may be combined with their social media data to suggest potentially most useful weak ties. This kind of combination of different kinds of data would provide a better understanding of participants' research interests and capabilities, resulting in more valuable recommendations. Lastly, event organizers should consider standardizing social media keywords for different discussion topics (e.g., use of specific for specific topics) across different social media channels. It will be helpful for the event organizers to find out what is the most relevant topic for the event as a whole. For event participants, it will help them locate potential partners to collaborate with or network with.

The use of potentially useful implicit social ties identified using existing social data, in the context of a business phenomena or business decision making will have some practical implications. In Experiment: Kickstarter, the role of implicit social ties (derived using social data) in the context of crowdfunding success has some practical implications towards the project owners and also the crowdfunding campaign managers. The results of the first part of the study show that at the broader level implicit social ties have an impact on crowdfunding success, while the results of the second part of the study specifically show that the social ties of the project owner have an impact on the degree of success of crowdfunding project. The results of Experiment: Kickstarter imply that the project owners need to directly interact as much as possible with other users on the social media platforms. At the same time, the project owners need to make an conscious effort to interact with and bring together users from disparate groups into the same conversation thread. These practices can increase the chances of bringing more funding for the crowdfunding project. At the same time, the way the social media data was analysed in the Experiment: Kickstarter provides an opportunity for the crowdfunding campaign managers to use similar analysis methods to gain insights about the implicit social ties based on the social media data. These insights can be used to reach more potential project backers and enable getting more funding for the crowdfunding project.

6.3 Evaluation of the study

This dissertation primarily used quantitative methodology. The most important criteria for evaluation of quantitative studies are reliability and validity (Bryman, Becker and Sempik 2008; Yilmaz 2013). Reliability means consistency or the degree to which the data collection methods and analysis procedures result in consistent findings (Saunders and Thornhill 2019; Yilmaz 2013). In other words, reliability means that if the research were repeated in the same context and used the same meth-

Table 6.3 Potential sources of biases in the research, mitigation of these biases and their impact on the conclusions

Sub-study	Potential main sources of bias in study	How these potential biases were avoided or addressed ?	Impact on conclusions / generalization
Case: Indiegame Developer Community	• Incomplete or missing data for the Facebook group related to the Indiegame developer community. (Potential impact on the reliability and validity of the study.)	• The completeness of the data collection was ensured by using the tool SODATO. This tool has been specifically developed to collect the data from open Facebook groups and pages. SODATO has been used to collect data for multiple research studies.	• The potential biases in the study were mitigated to a large extent. These potential biases do not have any major impact on the results of the study. The reliability and validity of the study was ensured by using appropriate data collection tool and using the established technique of social network analysis to develop the method to identify implicit social ties from social and big social data. Thus, the conclusions are drawn from these results can also be generalized to other relevant contexts related to the identification of implicit social ties from social and big social data.
Case: CMAD	• Incorrect or incomplete set of keywords were used for collecting the data related to the conference from Twitter REST API. (Potential impact on reliability & validity of the study) • Incomplete or missing data related to the conference on Facebook. (Potential impact on reliability & validity of the study)	• The Twitter data from Twitter REST API was collected using all the official hashtags (keywords) recommended by the conference organizers to event participants. • The Facebook data was collected from all the official Facebook groups and pages. This information was also cross-verified with the CMAD organizers to ensure that all relevant related to the event was collected. • The data from Facebook pages and groups were collected using the tool SODATO which has been used to collect Facebook data for multiple research projects. • The Twitter data was also cross-verified by comparing it with the Twitter data which was collected using Flockler – a tool used by the conference organizers for analyzing the Twitter data for CMAD. This enabled ensuring the completeness of the dataset • The social media usage analysis of the CMAD participants was also done to ensure that the results of the study did not have biases that could impact the internal validity of the study.	• The potential biases of the study were mitigated to a sufficient extent. The reliability and the validity of this study were ensured by using the original sources of data during the data collection process. Existing literature was used to develop measures that were used to carry out the data analysis. Also, established techniques like network analysis were used for analyzing the data. Thus, there was no significant impact on the results of the study. The conclusions drawn from the results of the study can be generalized to other similar knowledge work contexts.
Experiment: Kickstarter	• Using an incomplete or missing set of keywords while collecting data from Twitter. (Potential impact on the reliability of the study) • Having a biased data set related to only one kind of crowdfunding project. (Potential impact on the validity of the study)	• All the different combinations of the relevant keywords were used by also considering the case-sensitive nature of the keywords. For example, different upper and lower case combinations of the keywords were considered like Kickstarter, kickstarter, and other possible combinations while collecting data from Twitter API. • The different researchers involved in this study were asked to independently decide the different sets of keywords relevant for collecting the crowdfunding related data from Twitter. This was done to avoid the researcher's personal bias related to the relevant keywords for Twitter data collection. • The Twitter premium API was used for collecting the Twitter data to ensure the maximum possible completeness of the dataset. • The entire Kickstarter project data was collected to avoid the skewness of the crowdfunding project-related data.	• By adopting different measures, most of the potential biases of the study were avoided to a sufficient extent. The reliability of the study was ensured by using established statistical methods like regression analysis and ensuring the completeness of the data to a sufficient extent. The statistical significance of the results and the effect of control variables on the study along with the calculation of VIF values was done to ensure the validity of the study. Hence, the results are based on the analysis of the data do not contain any specific biases and the conclusions drawn from these results can be replicated and generalized to other relevant studies.

ods and the same data sources, similar results would be obtained (Shenton 2004) by some- one other than the researcher (Gummesson 2006; R. K. Yin 2018). Validity can be defined as whether the measures correspond closely to reality (Mark Easterby-Smith and Lowe 2003). In other words, validity can be defined as the extent to which a concept is accurately measured in a quantitative study (Heale and Twycross 2015). Validity can be further divided into internal validity and external validity. Internal validity, seeks to ensure that studies measure and test what is actually intended (Hardy and Bryman 2012; Shenton 2004). On the other hand, external validity is related to the extent to which the results of a particular study can be generalized to other relevant contexts beyond the present study, that is, other settings, people, programs, places,times, cases or approaches (Saunders and Thornhill 2019; Yilmaz 2013). Table 6.3 lists the major potential biases which the various sub-studies of this dissertation could have related to the reliability and validity. It also provides the different ways in which these potential biases were mitigated and also addresses how these potential biases impact the conclusions of the various sub-studies of this dissertation.

Reliability: In order for the research to be reliable and valid, the data used should be of high quality (Bryman, Becker and Sempik 2008; Yilmaz 2013). In this dissertation, the quality of the data was ensured by obtaining the data from official or original sources. In Case: CMAD, the social media related to CMAD2016 was collected using the Twitter REST API in case of Twitter and using the tool SODATO (Hussain and Vatrapu 2014) in case of Facebook. In Case: Indie Game Developer Community, the Facebook data related to the Indie Game Developer Community was collected using the tool SODATO (Hussain, Vatrapu et al. 2014). In Experiment: Kickstarter, the Twitter data related to the conversations of the crowdfunding project owners and potential project backers was collected using the Twitter Premium API. The data related to the crowdfunding projects for Kickstarter.com was accessed from the publicly available datasets for Kickstarter.com projects on the website webrobots.io. All the data handling processes were done carefully to ensure and maintain the high quality of the data.

Internal Validity: Furthermore, this research was conducted using the widely established statistical methods. In Case: CMAD, both questionnaire data and visual analytics were used to identify different kinds of implicit social ties in an event. In

Case: Indie Game Developer Community, both descriptive analytics and visual analytics were used to identify the potentially relevant and useful implicit social ties from the Indie Game Developer Community. Both the case studies used established mathematical analysis method of network analysis. Regression analysis was used in Experiment: Kickstarter to study the effect of implicit social ties on crowdfunding success. In both these studies, the effect of control variables on result was checked. The measures used in the different studies of this dissertation were carefully selected after reviewing prior literature to ensure that they measure the subjects of interest.

External Validity: The approach developed to identify potentially relevant and useful implicit social ties from social data in Case: Indie Game Developer Community was based on network analysis. Thus, this approach can be used in other relevant research settings as well and is not restricted to just a specific game development community or social media platform like Facebook. Hence, the approach developed in the Case: Indie Game Developer Community can be used in other contexts for identification of implicit social ties from social and big social data. A decision was made in Case: CMAD to focus on one extreme case - CMAD2016 to identify different implicit social ties from social data in the context of an event. However, the results of this study can be extended to other similar knowledge work context like professional conferences, research conferences and other similar contexts. For example, the study shows that there is a need to understand the use of a particular social media channel before using its data for implicit social tie identification. This result is relevant for other studies which use social media data for identification of implicit social ties in context of events. For the findings of the study to be applicable/ relevant for other similar contexts, the precondition that the social media is used by the event participants for networking and maintaining relationships should be fulfilled. Experiment: Kickstarter uses crowdfunding project data from the largest crowdfunding platform in the world. Hence, the sample is not representative of all the different kinds of crowdfunding platforms, but it can be considered as a representative of a reward-based crowdfunding platform which uses the "All-Or-Nothing" (AON) project funding format as described in 2.4.2.2. Thus, the results of this study can be applied or generalized to other reward-based crowdfunding platforms which follow the AON funding format.

6.4 Limitations of the study and future directions

The aim of this study was to open new ways to identify implicit social ties from social data and big social data and to study the role of the implicit social ties in different contexts. Many of the limitations of this thesis are related to the data.

While conducting the research study related to Case: CMAD, a limited number of survey responses were received. This limited our ability to draw any statistically significant results. Also these studies could only use some possible measures from social media data to identify different kinds of social ties. Though both these studies used two different social media platforms - Facebook and Twitter, the total size of this dataset was relatively small. In future, research can be done on more professional events like event and beyond to gather large number of survey responses in order to draw statistically significant results to further increase the accuracy and generalizability of results. Also newer possible measures from social media data can be developed to identify different kinds of social ties. In addition, effort can be made in future to incorporate larger datasets from social media and also different kinds of social data and big social data (like bibliographic data, location data, conference registration data) to identify different kinds of social ties.

In Case: Indie Game Developer Community, an edge-centered approach was used to develop a reciprocal interaction network. However, this study did not analyze the actual textual content of the interaction data. In future the different content analysis and text mining techniques can be used in combination with the edge-centered approach on large social data datasets.

In Experiment: Kickstarter the role of implicit social ties towards the crowdfunding project success was studied. In this sub-study, the R-squared value was found to be low. In general, a low R-squared value means that not a lot of the variation in the dependent variable is explained by the model. Even with a low R-squared value, statistically, significant coefficients continue to represent the mean change in the dependent variable given a one-unit shift in the independent variable. However, in the sub-study, some of the independent variables were found to be statistically significant and were useful in drawing important conclusions about the relationships between the variables. In future studies more independent variables could be included and also other analysis models could be used which could lead to a higher R-squared value and account for explaining to a higher degree the variation in the dependent variable by the model. Another limitation of this sub-study was that the experiments used social media data from only Twitter and the data related to crowdfunding projects was also limited to only one crowdfunding platform - Kickstarter.com. In the future, similar studies could be done using data from multiple social media platforms and also use crowdfunding data from multiple crowdfunding platforms.

At a broader level, most of the limitations of this dissertation are related to two main factors. The first factor was related to the limited amount of big social data and social data used for carrying out the dissertation. The second factor was related to the different contexts in which the different sub-studies were carried out. In future studies, there should be an emphasis on using a large variety of social and big social data to identify implicit social ties. At the same time, future studies should also focus on different contexts to identify implicit social ties from social and big social data. Future studies should also focus on understanding the role implicit social ties play in other varied contexts.

REFERENCES

Agrawal, A., Catalini, C. and Goldfarb, A. (Feb. 2011). *The Geography of Crowdfunding*. Tech. rep. Cambridge, MA: National Bureau of Economic Research. DOI: 10.3386/w16820.

— (June 2015). Crowdfunding: Geography, Social Networks, and the Timing of Investment Decisions. *Journal of Economics and Management Strategy* 24.2, 253–274. ISSN: 15309134. DOI: 10.1111/jems.12093.

Ahmed, A. M., Qiu, T., Xia, F., Jedari, B. and Abolfazli, S. (2014). Event-Based Mobile Social Networks: Services, Technologies, and Applications. *IEEE Access* 2, 500–513. ISSN: 2169-3536 VO - 2. DOI: 10.1109/ACCESS.2014.2319823.

Ahn, H. and Park, J.-H. (June 2015). The structural effects of sharing function on Twitter networks: Focusing on the retweet function. *Journal of Information Science* 41.3, 354–365. ISSN: 0165-5515. DOI: 10.1177/0165551515574974.

Alba, R. D. and Kadushin, C. (Aug. 1976). The intersection of social circles: A New Measure of Social Proximity in Networks. *Sociological Methods & Research* 5.1, 77–102. ISSN: 15528294. DOI: 10.1177/004912417600500103.

Aral, S. and Walker, D. (June 2014). Tie Strength, Embeddedness, and Social Influence: A Large-Scale Networked Experiment. *Management Science* 60.6, 1352–1370. ISSN: 0025-1909. DOI: 10.1287/mnsc.2014.1936.

Aramo-Immonen, H., Jussila, J. and Huhtamäki, J. (Oct. 2015). Exploring co-learning behavior of conference participants with visual network analysis of Twitter data. *Computers in Human Behavior*. Computing for Human Learning, Behaviour and Collaboration in the Social and Mobile Networks Era 51, 1154–1162. ISSN: 07475632. DOI: 10.1016/j.chb.2015.02.033.

Aramo-Immonen, H., Kärkkäinen, H., Jussila, J. J., Joel-Edgar, S. and Huhtamäki, J. (Feb. 2016). Visualizing informal learning behavior from conference participants' Twitter data with the Ostinato Model. *Computers in Human Behavior* 55, 584–595. ISSN: 07475632. DOI: 10.1016/j.chb.2015.09.043.

Backstrom, L. and Kleinberg, J. (2013). Romantic Partnerships and the Dispersion of Social Ties: A Network Analysis of Relationship Status on Facebook. *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing - CSCW '14*, 831–841. ISSN: 9781450325400. DOI: 10.1145/2531602.2531642.

Bakshy, E., Eckles, D., Yan, R. and Rosenn, I. (2012). Social influence in social advertising: Evidence from field experiments. *Proceedings of the ACM Conference on Electronic Commerce*. EC '12. New York, NY, USA: Association for Computing Machinery, 146–161. ISBN: 9781450314152. DOI: 10.1145/2229012.2229027.

Bastian, M., Heymann, S. and Jacomy, M. (2009). Gephi: An Open Source Software for Exploring and Manipulating Networks. *Third International AAAI Conference on Weblogs and Social Media*, 361–362. ISSN: 14753898. DOI: 10.1136/qshc.2004.010033.

Belleflamme, P., Lambert, T. and Schwienbacher, A. (Sept. 2014). Crowdfunding: Tapping the right crowd. *Journal of Business Venturing* 29.5, 585–609. ISSN: 08839026. DOI: 10.1016/j.jbusvent.2013.07.003.

Blondel, V. D., Guillaume, J. L., Lambiotte, R. and Lefebvre, E. (Oct. 2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* 2008.10, P10008. ISSN: 17425468. DOI: 10.1088/1742-5468/2008/10/P10008.

Blumstein, P. and Kollock, P. (Aug. 1988). Personal Relationships. *Annual Review of Sociology* 14.1, 467–490. ISSN: 0360-0572. DOI: 10.1146/annurev.so.14.080188.002343.

Borgatti, S. P., Everett, M. G. and Johnson, J. C. (2013). *Analyzing social networks*. English. Ed. by S. P. Borgatti, M. G. Everett and J. C. Johnson. SAGE Publications Inc. ISBN: 9781446247402 1446247406 9781446247419 1446247414.

Borgatti, S. P. and Halgin, D. S. (2011). On Network Theory. *Organization Science* 22.5, 1168. DOI: 10.1287/orsc.1100.0641.

Borst, I., Moser, C. and Ferguson, J. (2018). From friendfunding to crowdfunding: Relevance of relationships, social media, and platform activities to crowdfunding performance. *New Media and Society* 20.4, 1396–1414. ISSN: 14617315. DOI: 10.1177/1461444817694599.

Boyd, D., Golder, S. and Lotan, G. (2010). Tweet, tweet, retweet: Conversational aspects of retweeting on twitter. *Proceedings of the Annual Hawaii International*

Conference on System Sciences. ISBN: 9780769538693. DOI: 10.1109/HICSS.2010.412.

Bryman, A., Becker, S. and Sempik, J. (Oct. 2008). Quality criteria for quantitative, qualitative and mixed methods research: A view from social policy. *International Journal of Social Research Methodology* 11.4, 261–276. ISSN: 13645579. DOI: 10.1080/13645570701401644.

Burt, R. S. (June 1997). The Contingent Value of Social Capital. *Administrative Science Quarterly* 42.2, 339–365. ISSN: 00018392. DOI: 10.2307/2393923.

— (2001). *STRUCTURAL HOLES vs NETWORK CLOSURE*. Tech. rep., 31–56.

— (1992). *Structural Holes*. Harvard University Press. DOI: 10.4159/9780674029095.

— (Sept. 2004). Structural Holes and Good Ideas. *American Journal of Sociology* 110.2, 349–399. ISSN: 0002-9602. DOI: 10.1086/421787.

Burtch, G., Ghose, A. and Wattal, S. (June 2014). Cultural differences and geography as determinants of online prosocial lending. *MIS Quarterly: Management Information Systems* 38.3, 773–794. ISSN: 21629730. DOI: 10.25300/MISQ/2014/38.3.07.

Chen, D., Lü, L., Shang, M. S., Zhang, Y. C. and Zhou, T. (2012). Identifying influential nodes in complex networks. *Physica A: Statistical Mechanics and its Applications* 391.4, 1777–1787. ISSN: 03784371. DOI: 10.1016/j.physa.2011.09.017.

Chung, Y., Li, Y. and Jia, J. (Apr. 2021). Exploring embeddedness, centrality, and social influence on backer behavior: the role of backer networks in crowdfunding. *Journal of the Academy of Marketing Science*, 1–22. ISSN: 15527824. DOI: 10.1007/s11747-021-00779-x.

Colombo, M. G., Franzoni, C. and Rossi-Lamastra, C. (Jan. 2015). Internal social capital and the attraction of early contributions in crowdfunding. *Entrepreneurship: Theory and Practice* 39.1, 75–100. ISSN: 15406520. DOI: 10.1111/etap.12118.

Crawford, K. and Finn, M. (Aug. 2015). The limits of crisis data: analytical and ethical challenges of using social and mobile data to understand disasters. *GeoJournal* 80.4, 491–502. ISSN: 15729893. DOI: 10.1007/s10708-014-9597-z.

Creswell, J. W. and Creswell, J. D. (2017). *Research Design : Qualitative, Quantitative, and Mixed Methods Approaches*. English. 5th ed. Thousand Oaks : SAGE Publications. ISBN: 9781506386713 1506386717 9781506386683 1506386687.

Cumming, D. J., Leboeuf, G. and Schwienbacher, A. (June 2020). Crowdfunding models: Keep-It-All vs. All-Or-Nothing. *Financial Management* 49.2, 331–360. ISSN: 1755053X. DOI: 10.1111/fima.12262.

Dahlander, L. and McFarland, D. A. (Mar. 2013). Ties that last: Tie formation and persistence in research collaborations over time. *Administrative Science Quarterly* 58.1, 69–110. ISSN: 19303815. DOI: 10.1177/0001839212474272.

Daly, E. M. and Haahr, M. (May 2009). Social network analysis for information flow in disconnected delay-tolerant MANETs. *IEEE Transactions on Mobile Computing* 8.5, 606–621. ISSN: 15361233. DOI: 10.1109/TMC.2008.161.

, d. m. boyd danah m. and Ellison, N. B. (Oct. 2007). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication* 13.1, 210–230. ISSN: 10836101. DOI: 10.1111/j.1083-6101.2007.00393.x.

Davies, W. E. and Giovannetti, E. (Aug. 2018). Signalling experience & reciprocity to temper asymmetric information in crowdfunding evidence from 10,000 projects. *Technological Forecasting and Social Change* 133, 118–131. ISSN: 00401625. DOI: 10.1016/j.techfore.2018.03.011.

Davis, C. A., Ciampaglia, G. L., Aiello, L. M., Chung, K., Conover, M. D., Ferrara, E., Flammini, A., Fox, G. C., Gao, X., Gonçalves, B., Grabowicz, P. A., Hong, K., Hui, P. M., McCaulay, S., McKelvey, K., Meiss, M. R., Patil, S., Kankanamalage, C. P., Pentchev, V., Qiu, J., Ratkiewicz, J., Rudnick, A., Serrette, B., Shiralkar, P., Varol, O., Weng, L., Wu, T. L., Younge, A. J. and Menczer, F. (Oct. 2016). OSoMe: The IUNI observatory on social media. *PeerJ Computer Science* 2016.10, e87. ISSN: 23765992. DOI: 10.7717/peerj-cs.87.

De, L., Brass, D. J., Lu, Y. and Chen, D. (Sept. 2015). Friendships in online peer-to-peer lending: Pipes, prisms, and relational herding. *MIS Quarterly: Management Information Systems* 39.3, 729–742. ISSN: 21629730. DOI: 10.25300/misq/2015/39.3.11.

De Meo, P., Ferrara, E., Fiumara, G. and Provetti, A. (Oct. 2014). On facebook, most ties are weak. *Communications of the ACM* 57.11, 78–84. ISSN: 15577317. DOI: 10.1145/2629438.

Debortoli, S., Müller, O., Junglas, I. and Brocke, J. v. (July 2016). Text Mining For Information Systems Researchers: An Annotated Topic Modeling Tutorial. *Communications of the Association for Information Systems* 39.1, 7. ISSN: 1529-3181. DOI: 10.17705/1CAIS.03907.

Deri, S., Rappaz, J., Aiello, L. M. and Quercia, D. (Nov. 2018). Coloring in the links: Capturing social ties as they are perceived. *Proceedings of the ACM on Human-Computer Interaction* 2.CSCW, 1–18. ISSN: 25730142. DOI: 10.1145/3274312.

Dumais, S., Teevan, J., Jeffries, R., Russell, D. M. and Tang, D. (Jan. 2014). Understanding user behaviour through log data and analysis. *Ways of Knowing in HCI*. Springer New York, 350–372. ISBN: 9781493903788. DOI: 10.1007/978-1-4939-0378-8{_}14.

Easley, D. and Kleinberg, J. (July 2010). *Networks, Crowds, and Markets*. Cambridge University Press. ISBN: 9780521195331. DOI: 10.1017/cbo9780511761942.

Ebner, M., Mühlburger, H., Schaffert, S., Schiefner, M., Reinhardt, W. and Wheeler, S. (Sept. 2010). Getting granular on Twitter: Tweets from a conference and their limited usefulness for non-participants. *IFIP Advances in Information and Communication Technology*. Ed. by N. R. M. Turcsányi-Szabó. Vol. 324. Key Competencies in the Knowledge Society. Brisbane, Australia: Springer, 102–113. ISBN: 9783642153778. DOI: 10.1007/978-3-642-15378-5{_}10.

English, R. (Jan. 2014). Rent-a-crowd? Crowdfunding academic research. *First Monday* 19.1. ISSN: 13960466. DOI: 10.5210/fm.v19i1.4818.

Faraj, S., Kudaravalli, S. and Wasko, M. (June 2015). Leading collaboration in online communities. *MIS Quarterly: Management Information Systems* 39.2, 393–412. DOI: 10.25300/MISQ/2015/39.2.06.

Fogues, R. L., Such, J. M., Espinosa, A. and Garcia-Fornes, A. (Aug. 2018). Tie and tag: A study of tie strength and tags for photo sharing. *PLoS ONE* 13.8. Ed. by F. Xia, e0202540. ISSN: 19326203. DOI: 10.1371/journal.pone.0202540.

Fogués, R. L., Such, J. M., Espinosa, A. and Garcia-Fornes, A. (Sept. 2013). BFF: A tool for eliciting tie strength and user communities in social networking services. en. *Information Systems Frontiers* 16.2, 225–237. ISSN: 1387-3326, 1572-9419. DOI: 10.1007/s10796-013-9453-6.

Freedman, D. A. (2009). *Statistical models: Theory and practice*. English. Cambridge University Press. ISBN: 9780511815867. DOI: 10.1017/CBO9780511815867.

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks* 1.3, 215–239. ISSN: 03788733. DOI: 10.1016/0378-8733(78)90021-7.

Frydrych, D., Bock, A. J., Kinder, T. and Koeck, B. (2014). Exploring entrepreneurial legitimacy in reward-based crowdfunding. *Venture Capital* 16.3, 247–269. ISSN: 14645343. DOI: 10.1080/13691066.2014.916512.

Gandomi, A. and Haider, M. (Apr. 2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management* 35.2, 137–144. ISSN: 0268-4012. DOI: 10.1016/J.IJINFOMGT.2014.10.007.

Gerber, E. M. and Hui, J. (Dec. 2013). Crowdfunding: Motivations and deterrents for participation. *ACM Transactions on Computer-Human Interaction* 20.6, 1–32. ISSN: 10730516. DOI: 10.1145/2530540.

Geva, H., Oestreicher-Singer, G. and Saar-Tsechansky, M. (June 2019). Using retweets when shaping our online persona: Topic modeling approach. *MIS Quarterly: Management Information Systems* 43.2, 501–524. ISSN: 21629730. DOI: 10.25300/MISQ/2019/14346.

Ghauri, P., Grønhaug, K. and Strange, R. (2020). *Research Methods in Business Studies*. Cambridge University Press. ISBN: 9781108486743. DOI: 10.1017/9781108762427.

Gilbert, E. (2012). Predicting tie strength in a new medium. *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work - CSCW '12*. CSCW '12. New York, NY, USA: ACM, 1047. ISBN: 9781450310864. DOI: 10.1145/2145204.2145360.

Gilbert, E. and Karahalios, K. (2009). Predicting tie strength with social media. *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*. CHI '09. New York, NY, USA: ACM, 211. ISBN: 9781605582467. DOI: 10.1145/1518701.1518736.

Gilliland, D. I. and Bello, D. C. (Dec. 2002). Two sides to attitudinal commitment: The effect of calculative and loyalty commitment on enforcement mechanisms in distribution channels. *Journal of the Academy of Marketing Science* 30.1, 24–43. ISSN: 00920703. DOI: 10.1177/03079450094306.

Giudici, G., Nava, R., Rossi Lamastra, C. and Verecondo, C. (Oct. 2012). Crowdfunding: The New Frontier for Financing Entrepreneurship?: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.2157429.

Gleasure, R. and Feller, J. (2018). What kind of cause unites a crowd? Understanding crowdfunding as collective action. *Journal of Electronic Commerce Research* 19.3, 223–236. ISSN: 15266133.

Gouldner, A. W. (Apr. 1960). The Norm of Reciprocity: A Preliminary Statement. *American Sociological Review* 25.2, 161. DOI: 10.2307/2092623.

Granovetter (1973). The strength of weak ties. *American journal of sociology* 78.6, 1360. ISSN: 1939-1854.

Granovetter, M. (1983). The Strength of Weak Ties: A Network Theory Revisited. *Sociological Theory* 1, 201. ISSN: 07352751. DOI: 10.2307/202051.

Greiner, M. E. and Wang, H. (Jan. 2010). Building consumer-to-consumer trust in E-finance marketplaces: An empirical analysis. *International Journal of Electronic Commerce* 15.2, 105–136. ISSN: 10864415. DOI: 10.2753/JEC1086-4415150204.

Gummesson, E. (Jan. 2006). Qualitative research in management: addressing complexity, context and persona. *Management Decision* 44.2. Ed. by C. Cassell, 167–179. ISSN: 0025-1747. DOI: 10.1108/00251740610650175.

Gupta, J. P., Kärkkäinen, H., Li, H. and Mukkamala, R. R. (2022). Role of Implicit Social Ties in Crowdfunding Success. *Under Review*.

Gupta, J. P., Kärkkäinen, H., Torro, O. and Mukkamala, R. R. (2019). Revisiting social media tie strength in the era of data access restrictions. *IC3K 2019 - Proceedings of the 11th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*. Vol. 3. INSTICC. SciTePress, 187–194. ISBN: 9789897583827. DOI: 10.5220/0008067501870194.

Gupta, J. P., Li, H., Kärkkäinen, H. and Mukkamala, R. R. (2022). Social Ties' Role in Crowdfunding Projects' Success: A Project-Owner and Project-Backer Perspective. *Under Review*.

Gupta, J. P., Menon, K., Kärkkäinen, H., Huhtamäki, J., Mukkamala, R. R., Hussain, A., Vatrapu, R., Jussila, J. and Pirkkalainen, H. (2016). Identifying weak ties from publicly available social media data in an event. *AcademicMindtrek 2016 - Proceedings of the 20th International Academic Mindtrek Conference*. New York, New York, USA: ACM Press, 11–19. ISBN: 9781450343671. DOI: 10.1145/2994310.2994354.

Gupte, M. and Eliassi-Rad, T. (2012). Measuring tie strength in implicit social networks. *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12*. New York, New York, USA: ACM Press, 109–118. ISBN: 9781450312288. DOI: 10.1145/2380718.2380734.

Hansen, D. L., Shneiderman, B., Smith, M. A. and Himelboim, I. (Jan. 2020). Twitter: Information flows, influencers, and organic communities. *Analyzing Social*

Media Networks with NodeXL. Elsevier, 161–178. DOI: 10.1016/b978-0-12-817756-3.00011-x.

Hardy, M. and Bryman, A. (2012). Handbook of Data Analysis. *Handbook of Data Analysis*. DOI: 10.4135/9781848608184.

Harford, T. (Dec. 2014). *Big data: Are we making a big mistake*. DOI: 10.1111/j.1740-9713.2014.00778.x.

Haythornthwaite, C. (Oct. 2002). Strong, Weak, and Latent Ties and the Impact of New Media. *The Information Society* 18.5, 385–401. ISSN: 0197-2243. DOI: 10.1080/01972240290108195.

Heale, R. and Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-Based Nursing* 18.3, 66–67. ISSN: 14689618. DOI: 10.1136/eb-2015-102129.

Heide, J. B. and John, G. (Apr. 1992). Do Norms Matter in Marketing Relationships?: *Journal of Marketing* 56.2, 32. ISSN: 00222429. DOI: 10.2307/1252040.

Heider, F. (1958). *The psychology of interpersonal relations.* Hoboken, NJ, US: John Wiley & Sons Inc, ix, 326–ix, 326. DOI: 10.1037/10628-000.

Hervé, F., Manthé, E., Sannajust, A. and Schwienbacher, A. (May 2019). Determinants of individual investment decisions in investment-based crowdfunding. *Journal of Business Finance & Accounting* 46.5-6, 762–783. ISSN: 1468-5957. DOI: 10.1111/JBFA.12372.

Hobart, C., Perlman, D. and Duck, S. (22 1988). Intimate Relationships, Development, Dynamics and Deterioration. *Canadian Journal of Sociology / Cahiers canadiens de sociologie* 13.3, 329. ISSN: 03186431. DOI: 10.2307/3340860.

Hogan, B. (Dec. 2018). Social media giveth, social media taketh away: Facebook, friendships, and APIs. *International Journal of Communication* 12, 592–611. ISSN: 19328036.

Hong, Y., Hu, Y. and Burtch, G. (2018). Embeddedness, prosociality, and social influence: Evidence from online crowdfunding. *MIS Quarterly: Management Information Systems* 42.4, 1211–1224. ISSN: 21629730. DOI: 10.25300/MISQ/2018/14105.

Hornick, M. F. and Tamayo, P. (Aug. 2012). Extending recommender systems for disjoint user/item sets: The conference recommendation problem. *IEEE Transactions on Knowledge and Data Engineering* 24.8, 1478–1490. ISSN: 10414347. DOI: 10.1109/TKDE.2011.90.

Hornuf, L. and Schwienbacher, A. (Apr. 2017). Should securities regulation promote equity crowdfunding?: *Small Business Economics* 49.3, 579–593. ISSN: 15730913. DOI: 10.1007/s11187-017-9839-9.

Hosmer, D. W., Lemeshow, S. and Sturdivant, R. X. (2013). *Applied logistic regression.* English. Ed. by 3rd. ISBN: 9780470582473 0470582472.

Howison, J., Wiggins, A. and Crowston, K. (Dec. 2011). Validity issues in the use of social network analysis with digital trace data. *Journal of the Association for Information Systems* 12.12, 767–797. ISSN: 15369323. DOI: 10.17705/1jais.00282.

Huang, H., Tang, J., Liu, L., Luo, J. and Fu, X. (Dec. 2015). Triadic Closure Pattern Analysis and Prediction in Social Networks. *IEEE Transactions on Knowledge and Data Engineering* 27.12, 3374–3389. ISSN: 10414347. DOI: 10.1109/TKDE.2015.2453956.

Hui, J. S., Gerber, E. M. and Greenberg, M. (2012). Easy Money? The Demands of Crowdfunding Work. *Technical Report No. 4* 4, 1–11. URL: https://egerber.mech.northwestern.edu/wp-content/uploads/2012/11/Easy-Money-_The-Demands-of-CrowdfundingWork-_2012.pdf.

Hussain, A. and Vatrapu, R. (2014). Social Data Analytics Tool (SODATO). *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).* Vol. 8463 LNCS. Springer, Cham, 368–372. ISBN: 9783319067001. DOI: 10.1007/978-3-319-06701-8{_}27.

Hussain, A., Vatrapu, R., Hardt, D. and Jaffari, Z. A. (2014). Social Data Analytics Tool: A Demonstrative Case Study of Methodology and Software. *Analyzing Social Media Data and Web Networks*. London: Palgrave Macmillan UK, 99–118. DOI: 10.1057/9781137276773{_}5.

Hutto, C. J., Yardi, S. and Gilbert, E. (2013). A longitudinal study of follow predictors on Twitter. *Conference on Human Factors in Computing Systems - Proceedings*. New York, New York, USA: ACM Press, 821–830. ISBN: 9781450318990. DOI: 10.1145/2470654.2470771.

Ishikawa, H. (2015). *Social big data mining*. CRC Press, 1–261. ISBN: 9781498710947. DOI: 10.1201/b18223.

James, G., Witten, D., Hastie, T. and Tibishirani, R. (2013). *An Introduction to Statistical Learning with Applications in R*. Ed. by G. James, D. Witten, T. Hastie and

R. Tibishirani. Springer Publishing Company, Incorporated, 426. ISBN: 978-1-4614-7137-0.

Jones, J. J., Settle, J. E., Bond, R. M., Fariss, C. J., Marlow, C. and Fowler, J. H. (Jan. 2013). Inferring Tie Strength from Online Directed Behavior. *PLoS ONE* 8.1. Ed. by R. Lambiotte, e52168. ISSN: 1932-6203. DOI: `10.1371/journal.pone.0052168`.

Jussila, J., Huhtamäki, J., Kärkkäinen, H. and Still, K. (2013). Information visualization of Twitter data for co-organizing conferences. *Proceedings of the 17th International Academic MindTrek Conference: "Making Sense of Converging Media", MindTrek 2013*, 139–145. DOI: `10.1145/2523429.2523482`.

Kahanda, I. and Neville, J. (Mar. 2009). Using transactional information to predict link strength in online social networks. *Proceedings of the 3rd International AAAI Conference on Weblogs and Social Media*, 74–81.

Kang, L., Jiang, Q. and Tan, C. H. (Apr. 2017). Remarkable advocates: An investigation of geographic distance and social capital for crowdfunding. *Information and Management* 54.3, 336–348. ISSN: 03787206. DOI: `10.1016/j.im.2016.09.001`.

Kim, Y. and Zhang, Z. (Dec. 2018). Mobilizing Online Social Capital: the Relational View of Crowdfunding. *ICIS 2017: Transforming Society with Digital Innovation*. URL: `https://aisel.aisnet.org/icis2017/Peer-to-Peer/Presentations/2`.

Klöhn, L. (2018). The regulation of crowdfunding in Europe. *The Economics of Crowdfunding: Startups, Portals and Investor Behavior*. Ed. by D. Cumming and L. Hornuf. Cham: Springer International Publishing, 219–253. ISBN: 9783319661193. DOI: `10.1007/978-3-319-66119-3{\_}10`.

Koch, J. A. and Cheng, Q. (June 2016). The role of qualitative success factors in the analysis of crowdfunding success: Evidence from kickstarter. *Pacific Asia Conference on Information Systems, PACIS 2016 - Proceedings*. ISBN: 9789860491029.

Krackhardt, D. (1992). *The strength of strong ties: The importance of Philos in organizations*. DOI: `10.1177/1043463103154001`.

Krackhardt, D. and Stern, R. N. (June 1988). Informal Networks and Organizational Crises: An Experimental Simulation. *Social Psychology Quarterly* 51.2, 123–140. ISSN: 01902725. DOI: `10.2307/2786835`.

Kromidha, E. and Robson, P. (Oct. 2016). Social identity and signalling success factors in online crowdfunding. *Entrepreneurship and Regional Development* 28.9-10, 605–629. ISSN: 14645114. DOI: 10.1080/08985626.2016.1198425.

Kumar, V., Bhaskaran, V., Mirchandani, R. and Shah, M. (Jan. 2013). Creating a measurable social media marketing strategy: Increasing the value and ROI of intangibles and tangibles for Hokey Pokey. *Marketing Science* 32.2, 194–212. ISSN: 1526548X. DOI: 10.1287/mksc.1120.0768.

L. Fogues, R., Such, J. M., Espinosa, A. and Garcia-Fornes, A. (Aug. 2018). Tie and tag: A study of tie strength and tags for photo sharing. *PLOS ONE* 13.8. Ed. by F. Xia, e0202540. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0202540.

Lampe, C., Ellison, N. B. and Steinfield, C. (2008). Changes in Use and Perception of Facebook. *Proceedings of the ACM 2008 conference on Computer supported cooperative work - CSCW '08*. New York, New York, USA: ACM Press. ISBN: 9781605580074.

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A. L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D. and Van Alstyne, M. (Feb. 2009). *Social science: Computational social science*. DOI: 10.1126/science.1167742.

Levin, D. Z., Walter, J. and Murnighan, J. K. (June 2011). Dormant Ties: The Value Of Reconnecting. *Organization Science* 22.4, 923–939. ISSN: 10477039, 15265455.

Levin, D. Z. and Cross, R. (Nov. 2004). The Strength of Weak Ties You Can Trust: The Mediating Role of Trust in Effective Knowledge Transfer. *Management Science* 50.11, 1477–1490. ISSN: 0025-1909. DOI: 10.1287/mnsc.1030.0136.

Li, X., Cheng, X., Su, S., Li, S. and Yang, J. (2017). A hybrid collaborative filtering model for social influence prediction in event-based social networks. *Neurocomputing* 230, 197–209. ISSN: 0925-2312. DOI: https://doi.org/10.1016/j.neucom.2016.12.024.

Liberatore, F. and Quijano-Sanchez, L. (Sept. 2017). What do we really need to compute the Tie Strength? An empirical study applied to Social Networks. *Computer Communications* 110, 59–74. ISSN: 0140-3664. DOI: 10.1016/J.COMCOM.2017.06.001.

Lin, M. and Viswanathan, S. (May 2016). Home bias in online investments: An empirical study of an online crowdfunding market. *Management Science* 62.5, 1393–1414. ISSN: 15265501. DOI: 10.1287/mnsc.2015.2206.

Lin, N., Dayton, P. W. and Greenwald, P. (Nov. 1978). Analyzing the instrumental use of relations in the context of social structure. *Sociological Methods & Research* 7.2, 149–166. ISSN: 15528294. DOI: 10.1177/004912417800700203.

Lin, N., Ensel, W. M. and Vaughn, J. C. (June 1981). Social Resources and Strength of Ties: Structural Factors in Occupational Status Attainment. *American Sociological Review* 46.4, 393–405. ISSN: 00031224. DOI: 10.2307/2095260.

Lin, Y. and Boh, W. F. (Nov. 2020). How different Are crowdfunders? Examining archetypes of crowdfunders. *Journal of the Association for Information Science and Technology* 71.11, 1357–1370. ISSN: 23301643. DOI: 10.1002/asi.24332.

Liu, J. and Ding, J. (2020). Requesting for retweeting or donating? A research on how the fundraiser seeks help in the social charitable crowdfunding. *Physica A: Statistical Mechanics and its Applications* 557, 124812. ISSN: 03784371. DOI: 10.1016/j.physa.2020.124812.

Liu, S., Wang, B. and Xu, M. (2017). SERGE: Successive Event Recommendation Based on Graph Entropy for Event-Based Social Networks. *IEEE Access* 6, 3020–3030. ISSN: 21693536. DOI: 10.1109/ACCESS.2017.2786679.

Lu, C. T., Xie, S., Kong, X. and Yu, P. S. (2014). Inferring the impacts of social media on crowdfunding. *WSDM 2014 - Proceedings of the 7th ACM International Conference on Web Search and Data Mining*, 573–582. ISBN: 9781450323512. DOI: 10.1145/2556195.2556251.

Lu, Y., Qiao, Z., Zhou, C., Hu, Y. and Guo, L. (2016). Location-aware friend recommendation in event-based social networks: A Bayesian latent factor approach. *International Conference on Information and Knowledge Management, Proceedings*. Vol. 24-28-Octo. New York, New York, USA: ACM Press, 1957–1960. ISBN: 9781450340731. DOI: 10.1145/2983323.2983883.

Luarn, P. and Chiu, Y.-P. (Apr. 2015). Key variables to predict tie strength on social network sites. *Internet Research* 25.2, 218–238. ISSN: 1066-2243. DOI: 10.1108/IntR-11-2013-0231.

Madrazo-Lemarroy, P., Barajas-Portas, K. and Labastida Tovar, M. E. (Apr. 2019). Analyzing campaign's outcome in reward-based crowdfunding: Social capital as a determinant factor. *Internet Research*, 03–2018. ISSN: 10662243. DOI: 10.1108/INTR-03-2018-0115.

Mark Easterby-Smith, R. T. and Lowe, A. (2003). *Management Research: An Introduction*.

Marsden, P. V. and Campbell, K. E. (Dec. 1984). Measuring tie strength. *Social Forces* 63.2, 482–501. ISSN: 15347605. DOI: 10.1093/sf/63.2.482.

— (Sept. 2012). Reflections on conceptualizing and measuring tie strength. en. *Social Forces* 91.1, 17–23. ISSN: 00377732. DOI: 10.1093/sf/sos112.

Mathews, K. M., White, M. C., Long, R. G., Soper, B. and Von Bergen, C. W. (Sept. 1998). Association of indicators and predictors of tie strength. *Psychological Reports* 83.3 PART 2, 1459–1469. ISSN: 1558691X. DOI: 10.2466/pr0.1998.83.3f.1459.

Mattie, H., Engø-Monsen, K., Ling, R. and Onnela, J.-P. (Dec. 2018). Understanding tie strength in social networks using a local "bow tie" framework. *Scientific Reports* 8.1, 9349. ISSN: 2045-2322. DOI: 10.1038/s41598-018-27290-8.

Mcguire, G. M. and Bielby, W. T. (Aug. 2016). The Variable Effects of Tie Strength and Social Resources: How Type of Support Matters. *Work and Occupations* 43.1, 38–74. DOI: 10.1177/0730888415596560.

Mitchell, J. C. (Mar. 1987). The components of strong ties among homeless women. *Social Networks* 9.1, 37–47. ISSN: 03788733. DOI: 10.1016/0378-8733(87)90016-5.

Mollick, E. (Jan. 2014). The dynamics of crowdfunding: An exploratory study. *Journal of Business Venturing* 29.1, 1–16. ISSN: 08839026. DOI: 10.1016/j.jbusvent.2013.06.005.

Moritz, A. and Block, J. H. (2016). Crowdfunding: A literature review and research directions. *FGF Studies in Small Business and Entrepreneurship*. Springer, 25–53. DOI: 10.1007/978-3-319-18017-5{_}3.

Nitzan, I. and Libai, B. (Nov. 2011). Social effects on customer retention. *Journal of Marketing* 75.6, 24–38. ISSN: 00222429. DOI: 10.1509/jm.10.0209.

Noordewier, T. G., John, G. and Nevin, J. R. (Oct. 1990). Performance Outcomes of Purchasing Arrangements in Industrial Buyer-Vendor Relationships. *Journal of Marketing* 54.4, 80. ISSN: 00222429. DOI: 10.2307/1251761.

Olshannikova, E., Olsson, T., Huhtamäki, J. and Kärkkäinen, H. (Dec. 2017). Conceptualizing Big Social Data. *Journal of Big Data* 4.1, 1–19. ISSN: 21961115. DOI: 10.1186/s40537-017-0063-x.

Olteanu, A., Castillo, C., Diaz, F. and Kıcıman, E. (July 2019). Social Data: Biases, Methodological Pitfalls, and Ethical Boundaries. *Frontiers in Big Data* 2.13, 13. ISSN: 2624-909X. DOI: 10.3389/fdata.2019.00013.

Onnela, J.-P., Saramäki, J., Hyvönen, J., Szabó, G., Lazer, D., Kaski, K., Kertész, J. and Barabási, A.-L. (May 2007). Structure and tie strengths in mobile communication networks. *Proceedings of the National Academy of Sciences of the United States of America* 104.18, 7332–6. ISSN: 0027-8424. DOI: 10.1073/pnas.0610245104.

Oppermann, M. and Chon, K.-S. (1997). Convention participation decision-making process. *Annals of Tourism Research* 24.1, 178–191. ISSN: 0160-7383. DOI: https://doi.org/10.1016/S0160-7383(96)00049-7.

Ordanini, A., Miceli, L., Pizzetti, M. and Parasuraman, A. (Aug. 2011). Crowdfunding: Transforming customers into investors through innovative service platforms. *Journal of Service Management* 22.4, 443–470. ISSN: 17575818. DOI: 10.1108/09564231111155079.

Paschen, J. (Mar. 2017). Choose wisely: Crowdfunding through the stages of the startup life cycle. *Business Horizons* 60.2, 179–188. ISSN: 00076813. DOI: 10.1016/j.bushor.2016.11.003.

Patton, M. Q. (2002). *Qualitative research & evaluation methods : integrating theory and practice*, 806. ISBN: 9781412972123.

Pentland, A. (2015). *Social physics : how social networks can make us smarter*. English, 320. ISBN: 9780143126331.

Petersen, A. M. (Aug. 2015). Quantifying the impact of weak, strong, and super ties in scientific careers. *Proceedings of the National Academy of Sciences of the United States of America* 112.34, E4671–E4680. ISSN: 10916490. DOI: 10.1073/pnas.1501444112.

Petroczi, A., Bazsó, F. and Nepusz, T. (2007). Measuring tie-strength in virtual social networks. *Connections* 27.2, 39–52. ISSN: 02261766.

Pir Mohammadiani, R., Mohammadi, S. and Malik, Z. (Oct. 2017). Understanding the relationship strengths in users' activities, review helpfulness and influence. *Computers in Human Behavior* 75, 117–129. ISSN: 07475632. DOI: 10.1016/j.chb.2017.03.065.

Rapoport, A. (1953). Spread of information through a population with socio-structural bias: I. Assumption of transitivity. *The Bulletin of Mathematical Biophysics* 15.4, 523–533. ISSN: 00074985. DOI: 10.1007/BF02476440.

Reafee, W., Salim, N. and Khan, A. (May 2016). The Power of Implicit Social Relation in Rating Prediction of Social Recommender Systems. *PLOS ONE* 11.5. Ed. by Y. Wu, e0154848. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0154848.

Reinhardt, W., Ebner, M. and Beham, G. (2009). How people are using Twitter during conferences. *Creativity and Innovation Competencies on the Web, Hornung-Pr{ä}hauser, V., Luckmann, M. (Ed.) Proceeding of 5. EduMedia conference, p. 145-156, Salzburg*, 145–156. URL: http://www.edna.edu.au/edna/referral/browse/http://lamp.tu-graz.ac.at/~i203/ebner/publication/09_edumedia.pdf.

Reinhardt, W., Maicher, J., Drachsler, H. and Sloep, P. B. (2011). Awareness support in scientific event management with ginkgo. *Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies - i-KNOW '11*. New York, New York, USA: ACM Press, 1. ISBN: 9781450307321. DOI: 10.1145/2024288.2024336.

Retzer, S., Yoong, P. and Hooper, V. (2012). Inter-organisational knowledge transfer in social networks: A definition of intermediate ties. *Information Systems Frontiers* 14.2, 343–361. ISSN: 1572-9419. DOI: 10.1007/s10796-010-9250-4.

Ross, C., Terras, C., Warwick, M. and Welsh, A. (Mar. 2011). Enabled Backchannel: Conference Twitter use by Digital Humanists. *Journal of Documentation* 67.2, 214–237. ISSN: 0022-0418. DOI: 10.1108/00220411111109449.

Saunders Mark; Lewis, P. and Thornhill, A. (2019). *Research Methods for Business Students Eigth edition*. English. January, 2–30. ISBN: 978-0-273-71686-0.

Saxton, G. D. and Wang, L. (Oct. 2014). The Social Network Effect: The Determinants of Giving Through Social Media. *Nonprofit and Voluntary Sector Quarterly* 43.5, 850–868. ISSN: 08997640. DOI: 10.1177/0899764013485159.

Schuler, D. (Jan. 1994). Social computing. *Communications of the ACM* 37.1, 28–29. ISSN: 00010782.

Schwienbacher, A. (Dec. 2018). Entrepreneurial risk-taking in crowdfunding campaigns. *Small Business Economics* 51.4, 843–859. ISSN: 15730913. DOI: 10.1007/s11187-017-9965-4.

Servia-Rodríguez, S., Díaz-Redondo, R. P., Fernández-Vilas, A., Blanco-Fernández, Y. and Pazos-Arias, J. J. (Apr. 2014). A tie strength based model to socially-enhance applications and its enabling implementation: MySocialSphere. *Expert Systems with Applications* 41.5, 2582–2594. ISSN: 09574174. DOI: 10.1016/j.eswa.2013.10.006.

Severt, D., Wang, Y., Chen, P.-J. and Breiter, D. (2007). Examining the motivation, perceived performance, and behavioral intentions of convention attendees: Evidence from a regional conference. *Tourism Management* 28.2, 399–408. ISSN: 0261-5177. DOI: https://doi.org/10.1016/j.tourman.2006.04.003.

She, J., Tong, Y., Chen, L. and Cao, C. C. (Sept. 2016). Conflict-Aware Event-Participant Arrangement and Its Variant for Online Setting. *IEEE Transactions on Knowledge and Data Engineering* 28.9, 2281–2295. ISSN: 1041-4347. DOI: 10.1109/TKDE.2016.2565468.

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information* 22, 63–75. ISSN: 1875-8649. DOI: 10.3233/EFI-2004-22201.

Shneor, R. and Vik, A. A. (2020). Crowdfunding success: a systematic literature review 2010–2017. *Baltic Journal of Management* 15.2, 149–182. ISSN: 17465265. DOI: 10.1108/BJM-04-2019-0148. URL: https://www.emerald.com/insight/1746-5265.htm.

Sosa, M. E. (Feb. 2011). Where Do Creative Interactions Come From? The Role of Tie Content and Social Networks. *Organization Science* 22.1, 1–21. ISSN: 1047-7039. DOI: 10.1287/orsc.1090.0519.

Stanko, M. A., Bonner, J. M. and Calantone, R. J. (Nov. 2007). Building commitment in buyer-seller relationships: A tie strength perspective. *Industrial Marketing Management* 36.8, 1094–1103. ISSN: 00198501. DOI: 10.1016/j.indmarman.2006.10.001.

Stier, S., Breuer, J., Siegers, P. and Thorson, K. (Oct. 2020). *Integrating Survey Data and Digital Trace Data: Key Issues in Developing an Emerging Field*. DOI: 10.1177/0894439319843669.

Tan, Y. H. and Reddy, S. K. (Jan. 2021). Crowdfunding digital platforms: Backer networks and their impact on project outcomes. *Social Networks* 64, 158–172. ISSN: 03788733. DOI: 10.1016/j.socnet.2020.09.005.

Thies, F., Wessel, M. and Benlian, A. (Nov. 2018). Network effects on crowdfunding platforms: Exploring the implications of relaxing input control. *Information Systems Journal*. Vol. 28. 6. Blackwell Publishing Ltd, 1239–1262. DOI: 10.1111/isj.12194.

Tong, Y., She, J. and Meng, R. (2016). Bottleneck-aware arrangement over event-based social networks: the max-min approach. *World Wide Web* 19.6, 1151–1177. ISSN: 15731413. DOI: 10.1007/s11280-015-0377-6.

Torro, O., Gupta, J. P., Kärkkäinen, H., Pirkkalainen, H., Vatrapu, R. and Hussain, A. (2017). Integrating micro-level interactions with social network analysis in tie strength research: The edge-centered approach. *Proceedings of the 21st International Academic Mindtrek Conference, AcademicMindtrek 2017*. Vol. 2017-Janua. New York, New York, USA: ACM Press, 203–209. ISBN: 9781450354264. DOI: 10.1145/3131085.3131095. URL: http://dl.acm.org/citation.cfm?doid=3131085.3131095.

Tufekci, Z. (May 2014). *Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls*. Tech. rep. 1. URL: www.aaai.org.

Verlegh, P. W., Ryu, G., Tuk, M. A. and Feick, L. (Nov. 2013). Receiver responses to rewarded referrals: The motive inferences framework. *Journal of the Academy of Marketing Science* 41.6, 669–682. ISSN: 00920703. DOI: 10.1007/s11747-013-0327-8.

Villanueva-Felez, A., Woolley, R. and Cañibano, C. (Feb. 2015). Nanotechnology researchers' collaboration relationships: A gender analysis of access to scientific information. *Social Studies of Science* 45.1, 100–129. ISSN: 14603659. DOI: 10.1177/0306312714552347.

Walter, J., Levin, D. Z. and Murnighan, J. K. (Oct. 2015). Reconnection Choices: Selecting the Most Valuable (vs. Most Preferred) Dormant Ties. *Organization Science* 26.5, 1447–1465. ISSN: 1047-7039. DOI: 10.1287/orsc.2015.0996.

Walther, J. B. (June 1992). Interpersonal Effects in Computer-Mediated Interaction: A Relational Perspective. *Communication Research* 19.1, 52–90. ISSN: 15523810. DOI: 10.1177/009365092019001003. URL: https://journals.sagepub.com/doi/abs/10.1177/009365092019001003.

Wang, J. (Feb. 2016). Knowledge creation in collaboration networks: Effects of tie configuration. *Research Policy* 45.1, 68–80. ISSN: 00487333. DOI: 10.1016/j.respol.2015.09.003.

Wang, N., Liang, H., Xue, Y. and Ge, S. (2021). Mitigating information asymmetry to achieve crowdfunding success: Signaling and online communication. *Journal of the Association for Information Systems* 22.3, 773–796. ISSN: 15583457. DOI: 10.17705/1jais.00679.

Wegge, D., Vandebosch, H., Eggermont, S. and Walrave, M. (June 2015). The Strong, the Weak, and the Unbalanced The Link Between Tie Strength and Cyberaggression on a Social Network Site. en. *Social Science Computer Review* 33.3, 315–342. ISSN: 0894-4393, 1552-8286. DOI: 10.1177/0894439314546729.

Weller, K., Bruns, A., Burgess, J., Mahrt, M. and Puschmann, C. (2013). *Twitter and Society*. Bern, Switzerland: Peter Lang. ISBN: 978-1-4541-9992-2.

Wellman, B. (1982). *Studying personal communities in East York*. Vol. 128. Centre for Urban and Community Studies, University of Toronto Toronto ...

Wellman, B. and Wortley, S. (Nov. 1990). Different Strokes from Different Folks: Community Ties and Social Support. *American Journal of Sociology* 96.3, 558–588. ISSN: 0002-9602. DOI: 10.1086/229572.

White, R. W. (2013). Beliefs and Biases in Web Search. *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval*. New York, NY, USA: ACM. ISBN: 9781450320344.

Wiese, J., Min, J.-K., Hong, J. I. and Zimmerman, J. (2015). "You Never Call, You Never Write" *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15*. New York, New York, USA: ACM Press, 765–774. ISBN: 9781450329224. DOI: 10.1145/2675133.2675143.

Wu, A., DiMicco, J. M. and Millen, D. R. (2010). Detecting professional versus personal closeness using an enterprise social network site. *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*. New York, New York, USA: ACM Press, 1955. ISBN: 9781605589299. DOI: 10.1145/1753326.1753622.

Wuchty, S. (Sept. 2009). What is a social tie?: *Proceedings of the National Academy of Sciences* 106.36, 15099–15100. ISSN: 0027-8424. DOI: 10.1073/PNAS.0907905106.

Xia, F., Asabere, N. Y., Rodrigues, J. J. P. C., Basso, F., Deonauth, N. and Wang, W. (Dec. 2013). Socially-aware venue recommendation for conference participants. *Proceedings - IEEE 10th International Conference on Ubiquitous Intelligence and Computing, UIC 2013 and IEEE 10th International Conference on Autonomic and Trusted Computing, ATC 2013*. IEEE, 134–141. ISBN: 9781479924813. DOI: 10.1109/UIC-ATC.2013.81.

Xiang, R., Neville, J. and Rogati, M. (2010). Modeling relationship strength in online social networks. *Proceedings of the 19th international conference on World wide web - WWW '10*, 981. ISBN: 9781605587998. DOI: 10.1145/1772690.1772790.

Yi, Y., Zhang, Z. and Gan, C. (2018). The effect of social tie on information diffusion in complex networks. *Physica A: Statistical Mechanics and its Applications* 509, 783–794. ISSN: 03784371. DOI: 10.1016/j.physa.2018.06.063.

Yilmaz, K. (June 2013). Comparison of quantitative and qualitative research traditions: Epistemological, theoretical, and methodological differences. *European Journal of Education* 48.2, 311–325. ISSN: 01418211. DOI: 10.1111/ejed.12014.

Yin, C., Liu, L. and Mirkovski, K. (2019). Does more crowd participation bring more value to crowdfunding projects? The perspective of crowd capital. *Internet Research* 29.5, 1149–1170. ISSN: 10662243. DOI: 10.1108/INTR-03-2018-0103.

Yin, R. K. (2018). *Case study research and applications : design and methods*. English. SAGE Publications. ISBN: 9781506336169 1506336167 9781544308388 1544308388.

Yu, Z., Du, R., Guo, B., Xu, H., Gu, T., Wang, Z. and Zhang, D. (2015). Who should i invite for my party? combining user preference and influence maximization for social events. *UbiComp 2015 - Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. UbiComp '15. New York, NY, USA: Association for Computing Machinery, 879–883. ISBN: 9781450335744. DOI: 10.1145/2750858.2805839.

Zafarani, R., Abbasi, M. A. and Liu, H. (2014). *Social media mining: An introduction*. Vol. 9781107018. USA: Cambridge University Press, 1–320. ISBN: 9781139088510. DOI: 10.1017/CBO9781139088510.

Zhan Shi, Huaxia Rui and Whinston, A. B. (Mar. 2014). Content Sharing in a Social Broadcasting Environment: Evidence from Twitter. *MIS Quarterly* 38.1, 123–A6. ISSN: 02767783.

Zhang, A., Bhardwaj, A. and Karger, D. (2016). Confer: A Conference Recommendation and Meetup Tool. *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion - CSCW '16 Companion*. New York, New York, USA: ACM Press, 118–121. ISBN: 9781450339506. DOI: 10.1145/2818052.2874340.

Zhang, N., Datta, S. and Kannan, K. N. (Nov. 2014). An Analysis of Incentive Structures in Collaborative Economy: An Application to Crowdfunding Platform. *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.2518662.

Zhang, Y., Tan, C. D., Sun, J. and Yang, Z. (Nov. 2020). Why do people patronize donation-based crowdfunding platforms? An activity perspective of critical success factors. *Computers in Human Behavior* 112, 106470. ISSN: 0747-5632. DOI: 10.1016/J.CHB.2020.106470.

Zhao, J., Wu, J., Feng, X., Xiong, H. and Xu, K. (2012). Information propagation in online social networks: A tie-strength perspective. *Knowledge and Information Systems* 32.3, 589–608. ISSN: 02193116. DOI: 10.1007/s10115-011-0445-x.

Zheng, H., Li, D., Wu, J. and Xu, Y. (June 2014). The role of multidimensional social capital in crowdfunding: A comparative study in China and US. *Information and Management* 51.4, 488–496. ISSN: 03787206. DOI: 10.1016/j.im.2014.03.003.

Zhong, Y., Yang, J. and Nugroho, R. (June 2015). Incorporating Tie Strength in Robust Social Recommendation. *Proceedings - 2015 IEEE International Congress on Big Data, BigData Congress 2015*. IEEE, 63–70. ISBN: 9781467372787. DOI: 10.1109/BigDataCongress.2015.19.

Zhou, W., Duan, W. and Piramuthu, S. (Dec. 2014). A social network matrix for implicit and explicit social network plates. *Decision Support Systems* 68, 89–97. ISSN: 01679236. DOI: 10.1016/j.dss.2014.09.006.