

Software Metrics for Social Capital in Social Media

By Dawn E Carmichael

Doctor of Philosophy, 2015

Abstract

The aim of this research was creating metrics for measuring social connectedness in social media. This thesis made use of social capital theory in order to inform the construction of original metrics. The methodology used in this thesis involved conducting a literature review into the use of social capital theory in social media, proposing new metrics, implementation in software, validation, evaluation against other measures and finally demonstrating the utility of the new metrics.

A preliminary case study verified the suitability of using Facebook as a context for developing the metrics. The main practical work outlined in this thesis aimed to validate Social Capital in Social Media (SCiSM) metrics against the Internet Social Capital Scale (ISCS) (Williams 2006). The SCiSM metrics were developed to relate to bonding social capital, bridging social capital and total social capital (Putnam 2000). The methodology used to validate the SCiSM metrics was Meneely (2012) and involved using two independent data sets to validate the SCiSM metrics using both correlations and linear regression. Statistical analysis found a strong positive correlation between ISCS and SCiSM whilst regression analysis demonstrated that the relationship between SCiSM and ISCS was concerned with ranking rather than an absolute number. SCiSM was evaluated against other social capital metrics used in the literature such as degree centrality. It was found that SCiSM had a higher number of significant correlations with the ISCS than other measures.

The SCiSM metrics were then used to analyse the two independent data sets in order to demonstrate their utility. The first data set, taken from a Facebook group, was analysed using a paired t-test. It was found that bonding social capital increased over a twelve week period but that bridging social capital did not. The second data set, which was taken from Facebook status updates, was analysed using correlations. The result was that there was a positive

correlation between number of Facebook friends and bonding social capital. However it was also found that there was a negative correlation between number of Facebook friends and bridging social capital. This suggests that there is a dilution effect in the usefulness of large friend networks for bridging social capital.

In conclusion the problem that this research has addressed is providing a means to improve understanding of social capital in social media.

Permission to copy notice

Author: Dawn E Carmichael

Title: Software Metrics for Social Capital in Social Media

Qualification: Doctor of Philosophy

Date of Submission: 24th June 2015

A copy shall not be made of the whole or any part of the above mentioned thesis without the written consent of the undersigned.

Signature:

Address: University of Abertay Dundee, Bell Street, Dundee

Date: 18th October 2015

Acknowledgements

Many thanks in grateful acknowledgement of the advice and assistance from my supervisors Dr Jacqueline Archibald and Dr Geoffrey Lund. In addition, thanks for the assistance of Dr Matthew Craven who acted as a technical advisor on the mathematical aspects of this research project

Dedication

In memory of my darling mum Kay E Carmichael (1940-2008).

Table of Contents

Abstract	ii
Permission to copy notice	iv
Acknowledgements	v
Dedication	vi
Table of Contents	vii
Table of Figures	xii
Table of Tables	xiv
Table of Formulae	xv
Chapter 1 Introduction	1
1.1 Background	1
1.2 Rationale	1
1.3 Thesis research aims and objectives	5
1.4 Overview of Methodology	7
1.5 Thesis contribution	8
1.6 Thesis Structure	9
Chapter 2 Social Capital Theory	11
2.1 Overview	11
2.2 Introduction	11
2.3 Theories and Definitions of Social Capital	12
2.4 Critique of Social Capital Theory	20
2.5 Overview of Social Capital Theory	25

2.6	Theories of Social Capital used in Social Media	26
2.7	Critique of Social Capital in Social Media.....	37
2.8	Analysis of Key Contributors to the Literature	38
2.9	Overview of main points from the literature	40
2.10	Chapter Summary.....	41
Chapter 3 Measurement of Social Capital in Social Media.....		43
3.1	Overview	43
3.2	Social Media.....	43
3.3	Measuring Social Capital in Social Media Research	44
3.4	Social Network Analysis Measures in Social Media	61
3.4.1	Degree Centrality	63
3.4.2	Betweenness Centrality	64
3.4.3	Closeness Centrality.....	65
3.4.4	Clustering Coefficients.....	67
3.4.5	Eigenvector Centrality.....	68
3.5	Studies using Social Network Analysis in Social Media	70
3.6	Social Network Analysis Measures of Social Capital.....	72
3.7	Overview of Methods from the Literature	75
3.8	Overview of the Literature in Social Capital Measures in Social Media.....	77
3.9	Chapter Summary.....	79
Chapter 4 Preliminary Analysis of Social Media Usage		80
4.1	Overview	80

4.2	Preliminary Analysis: Student Facebook Groups	80
4.3	Interviews and survey.....	80
4.4	Social Network Analysis.....	83
4.5	Discussion	95
4.6	Ethical Research Procedures	96
4.7	Chapter Summary.....	98
Chapter 5 The SCiSM Metrics.....		100
5.1	Overview	100
5.2	Methodology for Validating the Metrics.....	101
5.3	Proposed Social Capital In Social Media (SCiSM) Metrics	103
5.3.1	Bonding SCiSM.....	104
5.3.2	Bridging SCiSM	107
5.3.3	Total SCiSM	110
5.3.4	SCiSM Formula Comparison to Formula in Literature	112
5.4	Implementing SCiSM in a re-usable extensible software framework.....	113
5.4.1	SCiSM Software Architecture	115
5.4.2	SCiSM Software & Web Site	116
5.5	Study 1: Validating the Proposed SCiSM Metrics.....	117
5.5.1	Aim of the study.....	118
5.5.2	Methodology.....	118

5.5.3	Results.....	119
5.5.3.1	Validation of proposed new metrics against ISCS Data Set 1	120
5.5.3.2	Validation of proposed new metrics against ISCS Data Set 2	124
5.5.3.3	Analysis of proposed SCiSM metrics against FIS	127
5.5.4	Discussion.....	129
5.6	Study 2: Evaluation of the SCiSM Metrics.....	131
5.6.1	Aim of the study.....	131
5.6.2	Methodology.....	131
5.6.3	Results.....	132
5.6.4	Discussion.....	133
5.7	Study 3 Using SCiSM to analyse social media: Facebook Group	135
5.8	Study 4 Using SCiSM to analyse social media Facebook status updates	139
5.9	Chapter Summary.....	142
Chapter 6 Conclusions		145
6.1	Introduction	145
6.2	Thesis Conclusions.....	145
6.3	Thesis Summary.....	146
6.4	Future Research.....	150
References.....		153
Appendix A Internet Social Capital Scales.....		162

Appendix B Preliminary Analysis Interview & Survey Questions	163
Appendix C SCiSM Software Code	170
Appendix D Data preparation for Data Set 1	177
Appendix E ISCS for Facebook Groups.....	180
Appendix F Facebook Intensity Scales.....	183
Appendix G Data Set 2 Facebook Status Updates Survey	184
Appendix H Data Preparation Data Set 2	187
Appendix I Data Sets 1 & 2	188
Appendix J Correlations evaluation of the measures for all participants	193

Table of Figures

Figure 3.1	Basic graph with node size given by degree centrality	63
Figure 3.2	Basic graph with node size given by betweenness score.....	65
Figure 3.3	Basic graph with node size given by closeness score	66
Figure 3.4	Basic graph with node size given by clustering coefficient.....	68
Figure 3.5	Basic graph with node size given by Eigenvector centrality score.....	70
Figure 4.1	Data Set A ‘degree centrality’ score indicated by node size	84
Figure 4.2	Data Set A Betweenness score indicated by node size (tutor top right)	86
Figure 4.3	Data Set B Betweenness score indicated by node size (tutor top right)	86
Figure 4.4	Data Set C Betweenness score indicated by node size (tutor top right)	87
Figure 4.5	Data Set A clustering coefficient score indicated by node size.....	88
Figure 4.6	Data Set B clustering coefficient score indicated by node size.....	89
Figure 4.7	Data Set C clustering coefficient score indicated by node size	89
Figure 4.8	Data Set A Pagerank (y-axis) Eigenvector (x-axis) degree centrality by size...91	
Figure 4.9	Data Set B Pagerank (y-axis) Eigenvector (x-axis) degree centrality by size ...92	
Figure 4.10	Data Set C Pagerank (y-axis) Eigenvector (x-axis) degree centrality by size92	
Figure 4.11	Data Set C Content Analysis by Degree Centrality	94
Figure 4.12	Data Set C Content Analysis by Eigenvector Centrality.....	94
Figure 5.1	Visualisation of a graph for an individual’s Bonding SCiSM.....	106
Figure 5.2	Visualisation of an individual’s Bridging SCiSM	108
Figure 5.3	Visualisation of graph for total SCiSM	111
Figure 5.4	Original MVC architecture	114
Figure 5.5	Revised MVC architecture used for SCiSM.....	115
Figure 5.6	SCiSM software architecture	116

Figure 5.7	SCiSM web site.....	117
Figure 5.8	Linear Regression bonding ISCS scores and SCiSM Metrics for data set 1 ..	122
Figure 5.9	Linear Regression Bridging ISCS scores and SCiSM Metrics for data set 1 ..	122
Figure 5.10	Linear Regression Total ISCS scores and SCiSM Metrics for data set 1	123
Figure 5.11	Linear Regression Total ISCS scores and SCiSM ranks for data set 1	124
Figure 5.12	Correlation between bonding ISCS scores and SCiSM Metrics data set 2..	125
Figure 5.13	Correlation between bridging ISCS scores and SCiSM Metrics data set 2 ..	126
Figure 5.14	Correlation between Total ISCS scores and SCiSM Metrics data set 2.....	126
Figure 5.15	Correlation of ranks of total ISCS scores and SCiSM Metrics data set 2....	127
Figure 5.16	Results of a t-test on total SCiSM scores for first vs second six weeks.....	136
Figure 5.17	Results of a t-test on SCiSM Bonding scores for first vs second six weeks	137
Figure 5.18	Results of a t-test on SCiSM Bridging scores for first vs second six weeks	137

Table of Tables

Table 1.1	Overview of Research Aims & Objectives	6
Table 3.1	Adjacency Matrix for nodes A,B and C from the kite network.....	69
Table 3.2	Adapted from Borgatti, Jones and Everett (1998)Mapping of SNA to SC	74
Table 3.3	Summary of Methods and Metrics for social capital from the literature.....	76
Table 4.1	Survey results of student use of the course Facebook group.....	81
Table 4.2	Survey results of student attitudes to the course Facebook group.....	82
Table 4.3	Degree Centrality Results	84
Table 4.4	Eigenvector centrality values across the 3 data sets	90
Table 5.1	Validity criteria mapped to research methods	102
Table 5.2	Correlation between ISCS and SCiSM Metrics.....	121
Table 5.3	Correlation between ISCS scores and SCiSM Metrics data set 2.....	125
Table 5.4	Correlation between FIS and SCiSM metrics.....	128
Table 5.5	Frequency of significant critical values for each score for all 8 participants ..	132
Table 5.6	Number of Facebook friends vs SCiSM scores correlations	140
Table 5.7	Number of Facebook posts vs SCiSM scores correlations	141

Table of Formulae

Formula 3.1	Group Density (Rafaeli, Ravid and Soroka 2004)	46
Formula 3.2	Individual Density (Rafaeli, Ravid and Soroka 2004)	46
Formula 3.3	Bonding Social Capital (Smith 2008)	52
Formula 3.4	Bridging social capital (Smith 2008)	53
Formula 3.5	Degree Centrality (Craven 2013)	64
Formula 3.6	Betweenness Centrality(Craven 2013).....	64
Formula 3.7	Closeness Centrality(Craven 2013).....	66
Formula 3.8	Clustering Coefficient (Craven 2013)	67
Formula 3.9	Eigenvector Centrality (Craven 2013)	69
Formula 5.1	Bonding SCiSM	105
Formula 5.2	Bridging SCiSM.....	108
Formula 5.3	Total SCiSM.....	110

Chapter 1 Introduction

1.1 Background

Social Media software has advanced over the last thirty-five years in accordance with the affordances of technology. From the early beginnings of communication systems on the Internet, for example the 'Bulletin Board Service' and 'The Well', where the technology enabled users to view and post messages in a discussion forum format. With the coming of 'Listserve' the use of email was transformed by allowing users to register to a group which received a common email message. The basic idea of grouping users has continued to be refined over time. The year 1999 saw the first of the modern Social Media giants emerged in the shape of 'blogger'. The year 2003 saw the launch of Facebook which has been a phenomenal commercial success and is by any measure the dominant Social Media web site at the present time. Today Facebook is cited either as the second most popular web site after Google (EBiz 2013) or the topped ranked site (Alexa Internet 2015). The Facebook Social Media web site is also dominant in terms of its market penetration, for instance in April 2015 the global user base for Facebook is thought to be 890 million daily active users (Checkfacebook 2015).

1.2 Rationale

Social media systems are extremely popular as evinced in popular culture on a daily basis. However this popularity disguises a concern that not everyone is able to take full advantage of the social affordances provided by the software. If a person has few social connections to begin with, the software may not perform as well for them as for more socially connected people. On the other hand, it might be the case that as individuals build up ever larger social networks the returns in terms of social value, i.e. social connectedness, may diminish. This uncertainty regarding the effectiveness of social media provided an impetus this research.

Despite, or perhaps even because of, the prodigious growth in the use of social media there is a lack of understanding of the social value that users obtain from social communication within these sites. The fact that there are so many users carrying out so many interactions means that the social connectedness derived from the software has been taken axiomatically. However it is far from self evident whether or not social value is equally available to all users regardless of their friendship networks and their usage patterns. This lack of understanding about the social value to users that social media provides means that the central purpose of the software is obscured. A clearer understanding about the users' experience of developing social connectedness using social media could potentially be beneficial in a number of different contexts. For instance examining social connectedness for specific user groups such as students in education could help to evaluate the software in particular settings. Understanding the social value experienced by users could also assist developers in evaluating new features in social media platforms. So to the use of social media in businesses and organisations is currently poorly understood (Sun and Shang 2014) and a means of examining the potential benefits would add to an analysis.

At the present time it is common in the various forms of social media analysis to use numeric measures of social connections such as; likes, follows and followers. However it has been suggested that these interactions provided skewed data as they are de-contextualised from a stream of interactions to a single data point (Baym 2013). The argument that there is skewing of 'likes' in social media can be boiled down to the notion that if one person has ten likes and another has five, the person with the highest number of 'likes' must be more positively valued. By extension into the arena of software metrics the assumption is that the software is more effective for the person with the higher number of 'likes'. The limitation of this de-contextualised data point approach is illustrated by the fact that, as of October 2014, Justin Bieber has 55,601,028 followers whilst Barak Obama has 47,979,937

(Twittercounter.com 2014). Leaving aside the intriguing socio-political connotations the number of twitter followers may have, it seems improbable that the effectiveness of the Twitter software can be derived from this type of raw number based statistic. In this sceptical vein regarding basic numeric measures, it has been suggested by Baym (2013), as well as Wang, Burke and Kraut (2013), that numeric social media metrics may lack consideration of the variations in audience and also the nature of the interactions which can be either positive and negative but that are not represented by simple counting of endorsements such as Facebook 'likes' and Twitter 'followers'.

In the search for valid metrics for social media it is worth noting that metrics are essentially created in order to examine an attribute of a system in greater detail. In this research the process of developing metrics involved creating a logical representation of a theoretical construct. As the attribute under examination is 'social value' i.e. social connectedness, social capital theory was investigated. The notion of social capital is an interesting concept, which provides a prima facie relevance to an exposition of software designed to be social in nature as social capital is purported to be a measure of the human assets that a person can access. The central idea of social capital originated in 1916 (Hanifan 1916) and gained popularity in the later decades of the twentieth century as a means for analysing civic engagement (Putnam 1993). In the intervening years social scientists and public policy researchers have made use of the idea of social capital because of its potential to provide a means to identify and explain social group phenomena. In relatively recent years the usage of the term social capital has spread to computer science due to the growth in online communities and social media (Ellison et al. 2014).

However despite years of research into social capital there is lack of valid metrics which can be implemented in software. The fundamental problem of social capital theory is the absence

of consensus on how to make a measurement (Fukuyama 2001). Furthermore the theoretical foundations for social capital metrics in computing are neither comprehensive nor rigorous (McCalla 2000). It has been suggested that social capital is poorly understood in the context of the web (Arvidsson 2011; Bermejo 2007). Despite these problems metrics for social capital hold out the possibility of providing a tool for analysis of social data which would otherwise be difficult to measure. The outcome of developing metrics of social capital is intended to be used to gain insight into the structural elements of group interactions, including knowledge sharing and social discourse in social media. Therefore it is suggested that a social capital metric, implemented in software, would enable researchers to investigate complex social constructs in a systematic fashion.

However there is a considerable gap between a complex sociological construct such as social capital theory and the measurement of effectiveness in social media software. To put this gap succinctly, to simply re-badge 'likes' as 'social capital' is not credible. There is no reason to suppose the number of 'likes' is related to social capital and still less to the social effectiveness of social media. The issues arising from the gap between social capital theory and metrics has in the past been addressed by Borgatti, Jones and Everett (1998) in a paper designed to map the mathematical techniques embodied in Social Network Analysis to social capital theory. This research seeks to update Borgatti's approach for the social media era.

In summary it is suggested there is a lack of valid metrics for social connectedness in social media and limitations in using user data such as 'likes' in measuring the performance of the software. Furthermore it is argued that a social capital metrics, based upon the techniques in Social Network Analysis, would make a useful contribution to the field of social media metrics. Therefore this thesis makes use of social capital theory in order to underpin metrics which can be outlined as a mathematical formula, validated and implemented in software.

1.3 Thesis research aims and objectives

The preceding sections suggested that there is a lack of valid metrics for social media and that social capital theory may provide the theoretical basis for the development of new metrics.

This section contains an explanation of what a software metric is and an outline of the methodology used in this research. This section also contains a description of the research aims and objectives.

The term ‘software metric’ has been succinctly defined as:

“a quantitative scale and method which can be used to determine the value a feature takes for a specific software product” (IEEE 1990)

In this vein it has also been suggested that a software metric is an attempt to measure an attribute of a system (Fenton and Neil 2000). In this research the attribute of the software under examination is social value, a complex idea which is why it is suggested that it requires a theoretical framework such as social capital theory. The use of social capital theory in this manner suggests another key definition namely the concept of ‘validity’. It has been argued that in the process of proposing new metrics, researchers have a burden of proof to meet in order to persuade the research community that the metrics are acceptable for its intended use. Furthermore it is essential to ensure that the validation criteria are not arbitrarily arrived at but rather the criteria used support an appropriate philosophy for the task at hand (Meneely 2012). In practice using the Meneely (2012) methodology involved assessing the key aims of the metrics and from these, deriving criteria which in turn were used to construct the methodology.

The aim of this research is to **create, validate and evaluate social capital metrics in order to analyse the social interaction performance of social media.**

The table 1.1 below provides an outline of the methodology in terms of the main research objectives, tasks and methods used in this research.

Research Objectives	Research Tasks	Research Methods
Examine social capital theory as it relates to social media	<ul style="list-style-type: none"> • Examine social capital theory and how it is used in social media. • Analyse how social capital has been measured in social media 	Literature review
Preliminary analysis of social capital in social media using a case study.	<ul style="list-style-type: none"> • Examine user attitudes to the case study software. • Examine the underlying properties of the social graph in the case study software. 	Interview /Survey Social Network Analysis
Develop metrics and related software, which are validated, evaluated and demonstrated.	<ul style="list-style-type: none"> • Propose social capital metrics. • Implement the metrics in a software framework. • Validate the proposed metrics against another valid measure of social capital using two independent data sets. • Evaluate how proposed metrics compare to other existing measures. • Demonstrate results yielded by using the new metrics using two data sets. 	Proposition Proof of concept Data and statistical analysis Data and statistical analysis Experimental and statistical analysis

Table 1.1 Overview of Research Aims & Objectives

An explanation of how the aim and objectives were met is provided in the next section.

1.4 Overview of Methodology

The literature review into social capital theory was a key part of this research as it provided the theoretical underpinning of the metrics. The literature review made reference to three components which were ‘social media metrics’, ‘Social Network Analysis’ (SNA) and ‘Social Capital Theory’. The focus of the literature lay in the intersection of these cognate areas as shown in figure 1.1.

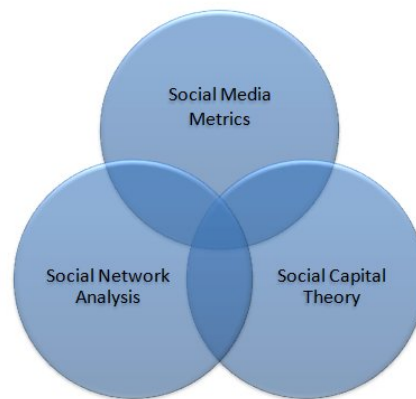


Figure 1.1 Intersection of Cognate Areas

The outcome of the literature review was the identification of fundamental aspects of what constitutes social capital and how it is currently measured in social media. The findings from the literature review are laid out in figure 1.2.

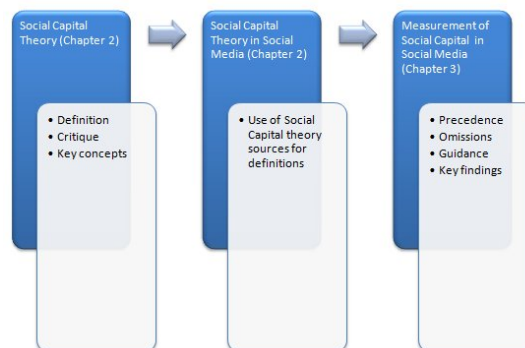


Figure 1.2 Organisation of Literature Review

The literature review was used to design the sequence of tasks making up the methodology as shown in Figure 1.3.

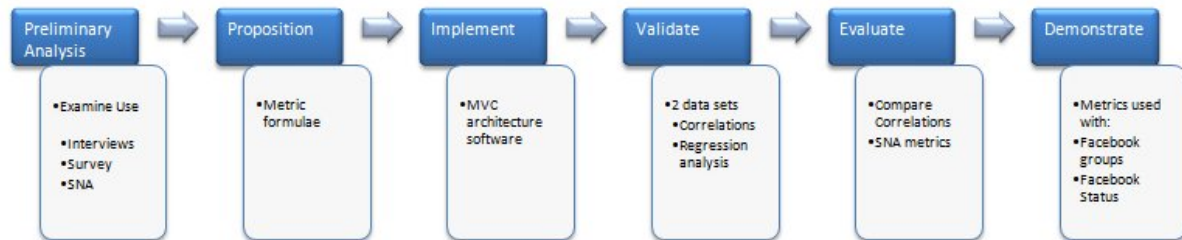


Figure 1.3 Overview of Methodology

As figure 1.3 indicates this research began with an initial preliminary study, which involved examining a case study of the use of social media to determine if it would make a suitable test bed. The next steps involved proposing Social Capital in Social Media (SCiSM) metrics in mathematical formulae and implementing them in software. Then the metrics were validated, using two new data sets, against self-reporting from users on their perceptions of their social capital. The validation was carried out using correlations and regression analysis. The next step was concerned with evaluating the proposed new metrics against other Social Network Analysis measures. The final stage was based upon demonstrating the utility of the metrics and examining the resultant data.

1.5 Thesis contribution

In summary the aim of this research is to contribute original metrics which can assist in analysing the performance of social media. In achieving the aim this thesis is built upon the contributions to the literature on social capital theory from sociologists such as Bourdieu (1985) and Putnam (2000). This research endeavours to expand the body of knowledge by using social capital theory and Social Network Analysis (SNA) in an innovative way. This contribution involves using social capital theory to create the Social Capital in Social Media (SCiSM) metrics. The SCiSM metrics are implemented in extensible software in order to

allow future researchers to analyse social media data. In addition this work contributes to research in field of measuring social capital by validating, evaluating and demonstrating the new social capital metrics.

1.6 Thesis Structure

The central aim of this research is 'Developing metrics for social connectedness in social media' and the process of fulfilling this aim is written in this thesis. The thesis is made up of six chapters which are outlined below.

The second and third chapter contains a literature review of social capital theory in three parts which are: 'social capital theory', 'social capital theory in the social media' and 'measurement of social capital in social media'. The second chapter contains the first two of the three parts. Amongst the findings in the chapter are the lack of a common definition for the term social capital and the potential usefulness of the ideas of bonding and bridging social capital, which relate to friends and acquaintances, respectively, in a social graph.

The third chapter contains the third part of the literature review, concerning how social capital theory is used in the study and measurement of social media. The chapter contains an explanation of how social capital is often measures by survey which includes the influential the Internet Social Capital Scale (ISCS) (Williams 2006). The chapter also contains the argument that Social Network Analysis measures represent a viable means of examining social capital in social media.

The fourth chapter contains the preliminary analysis of a case study of social media usage namely the use of a Facebook group in an academic setting. The preliminary analysis considered the main characteristics of the use of the software. The chapter concludes with

the suggestion that the case study is viable as a test bed for measuring social capital in social media.

The fifth chapter contains the proposed new social capital metrics and a study validating the metrics in terms of correlations and regression analysis with the Internet Social Capital Scale (ISCS) (Williams 2006) using two independent data sets. In addition the chapter contains a description of how the social capital metrics were implemented in software, using the Model View Controller architecture. The chapter also contains a study examining the performance of the proposed new social capital metrics against other SNA based measures such as degree centrality. The final two studies demonstrate the use of the software in examining social media.

The sixth chapter contains the conclusions of this research including suggested future work.

In summary the research work for this thesis includes the creation of new social capital metrics which are validated, evaluated, implemented in software and then used in order to demonstrate the potential of the metrics.

Chapter 2 Social Capital Theory

2.1 Overview

The literature review underpinning this thesis is comprised of three key components which are: 'social capital theory', 'social capital in social media' and the 'measurement of social capital in social media'. The literature review is given over two chapters. This chapter presents the first two components of the literature review the next chapter the third.

Therefore this chapter contains an examination of competing definitions of the term social capital and also provides a critique of social capital theory. In addition this chapter provides a discussion on how social capital theory has been used in the analysis of social media.

2.2 Introduction

“Capital is money, capital is commodities... By virtue of it being value, it has acquired the occult ability to add value to itself.” (Marx 1867)

Social capital theory arguably owes much to Karl Marx for making the word 'capital' more or less universally understood. At once upon reading the phrase 'social capital' the idea of social value can be conjured, whether derived by 'occult' means or otherwise. In the early part of the twentieth century social capital emerged as a concept used to understand the value of social networks. In the later part of the century social capital theory was used as a means of understanding social inequalities. Now in the early twenty first century the notion of 'social capital' is undergoing a new round of revision intended to make it useful in the realm of social media.

2.3 Theories and Definitions of Social Capital

A key feature of this research is to make use of theory to develop metrics for social media. To make use of theory in this way ensures that the metrics are transferable to different contexts. Therefore it is essential to review the definitions and central concepts in social capital theory.

One of the earliest mentions of social capital was made in an article in regard to schools in reference to engagement within rural communities (Hanifan 1916). The notion of social capital was differentiated from economic capital and described as being made up of:

“good will, fellowship, mutual sympathy and social intercourse”
(Hanifan 1916).

This collection of positive attributes, it was argued, can be directed towards the common good. However Hanifan makes the observation that in a rural community social capital may be in short supply. It was argued that in order to begin the process of community improvement, there must be sufficient community level social capital accrued. Social capital could be accrued; it was stated, through the use of local schools as venues for people to meet. For Hanifan, then, there is a concept of individual social capital which can be built up by the process of social interactions. The concept of social capital was also argued as being an attribute of a social grouping. The key ideas, of increasing social capital through interactions which require a venue as stated in Hanifan’s work of 1916 have a resonance with this research, although in this instance the venue is a virtual one.

In the more recent past social capital theory was revived by Granovetter (1973). Granovetter (1973) argued that social capital was simply the connections between individuals in a network and that these connections were heterogeneous. In specific terms it was argued that the heterogeneity was based on the social connections being either strong or weak ties. This

analysis of ties, set out by Granovetter, has had considerable impact in the literature as it has been argued that not all connections are of equal value in terms of social capital.

Furthermore Granovetter argued that weak ties in a person's social network were more likely to have information that the individual does not possess themselves nor is possessed by their strong ties and therefore were potentially more likely to yield social capital (Granovetter 1973). However Granovetter's ideas were indirectly challenged in another analysis of the nature of ties within a network by Loury (1977) who argued that dense ties in networks were a necessary pre-condition to the development of social capital (Loury 1977). This type of discrepancy on key ideas between theorists has continued to be common in the literature.

Widespread academic usage of the term 'social capital' did not come about until the later part of the twentieth century, particularly from the 1980's onwards. Arguably the first systematic use of the term was to refer to the human resources individuals possessed:

“the aggregate of the actual or potential resources which are linked to possession of a durable network of more or less institutionalized relationships of mutual acquaintance or recognition”(Bourdieu 1980, 1985).

Bourdieu's contribution is arguably one of the most significant as he introduced the term into sociological discussion. Bourdieu's discourse on the concept of social capital focuses on the benefits of social capital to individuals by participation in groups.

Bourdieu's ideas on social capital theory asserted that profits accrue from group membership, and that individuals make use of investment strategies to gain value from groups.

Bourdieu's definition makes clear that social capital has different forms and is analogous to economic capital. A clear link is drawn between an individual's social capital and their ability to access economic resources.

Bourdieu's definition of social capital also encompasses cultural capital which it was argued provides access to experts and to institutions, such as universities (Bourdieu 1985). A key argument put forward by Bourdieu (1985) deals with the interaction between money capital, social capital and cultural capital. The argument suggests that the various forms of capital are fungible, i.e., they can be traded for each other. Therefore social capital, it was argued, can seldom be acquired without the investment of economic and cultural resources required in order to create and maintain social relationships. Overall Bourdieu's treatise on social capital has been highly influential and is frequently cited in more recent contributions to the literature on social capital theory.

In the late 1980's Coleman defined social capital as relating to connections between entities with two common elements: shared social structure and the ability to facilitate action by actors. In Coleman's analysis actors can relate to individuals or organisations (Coleman 1988, 1990). Coleman appears to have used a rather vague definition of social capital, a tendency not restricted to him. The problem arises from including in the term a growing number of processes for acquiring social capital, some of which are unrelated to each other. The proliferation of process includes reciprocity of expectation, enforcement of norms and privileged access to information. However it is useful to distinguish resources from the ability to access them by virtue of group membership, as in fact Bourdieu does. Defining social capital as the resources acquired as Coleman does, rather than the enabling structures, can lead to over simplification. For example in the proposition that one person may have access to more financial support because they have higher levels of social capital than another does not take account of the fact that others may have relationships with people with less money supplying potential. Defining social capital in terms of the resources an individual may amass is potentially a limiting point of view and does not allow for the idea that the structural relationships implied by Bourdieu are worthy of analysis.

However Coleman does put forward a key concept in social capital theory namely asymmetry in the relationships, whereby some may be recipients of social capital and some may be donors (Coleman 1988, 1990). The motivation for recipients is fairly clear but is less clear for the donors. Therefore it was argued that an effective definition of social capital must include: recipients, sources and the resources. Overall Coleman's contribution has value in highlighting the importance of social capital in acquiring economic and cultural capital (Coleman 1988, 1990).

Coleman acknowledged Loury's influence in his work as well as a number of other economists; however, he does not comment on Bourdieu's work. This omission is striking as there are parallels in Coleman's idea of human capital and Bourdieu's cultural capital. Both of these ideas are concerned with acquiring social value by learning, in the case of Coleman: from role models, family and other environments, and from Bourdieu: via family, acquaintances and organisations.

Following on from the contributions made by Bourdieu, Loury and Coleman there a number of other theoreticians who have put forward definitions of social capital. An example is provided by Baker who defined social capital as a resource that social actors could obtain through social structures in order to derive benefit (Baker 1990). Another fairly broad definition was provided by Schiff who defined the term as elements of a social structure among individuals who are inputs to function (Schiff 1992). However a slightly more detailed definition of social capital was provided by Burt (1992), where the emphasis was put on friends and acquaintances from whom one receives opportunities to receive and use both social and financial capital (Burt 1992). Burt also asserted that the absence of ties, which he called structural holes, was valuable because dense networks conveyed large amounts of redundant information whilst networks with few ties could be more easily understood as

sources of new information and resources. As was the case with Granovetter (1973), for Burt a social network with sparse ties is preferable to one with dense ties because it provides a more effective means of transmitting social capital. This argument is at odds with Coleman and Loury who argued that dense networks were a necessary precondition for the advancement of social capital. These opposing points of view are especially important to this thesis because both of these conditions, dense and sparse, can be prevalent in social media networks.

Another highly influential contribution to social capital theory suggests that the term refers to social networks and their associated norms of reciprocity (Putnam 1993, 1995, 2000).

Putnam puts the emphasis on social capital being related to both the network and the effect of the network. Other significant ideas from Putnam include the notions of trust in networks and the importance of reciprocity. Putnam's contribution defines social capital as being made of a network of social connections which can be divided into bonding and bridging social capital. Bonding social capital refers to the types of sustaining interactions typical of friends and families, whilst bridging social capital refers to the type of interactions with acquaintances which are used to help achieve objectives such as information discovery (Putnam 2000). Putnam's work on bonding and bridging social capital is particularly important because the concepts articulate with greater clarity earlier work on the nature of ties in a network. Earlier contributions used the ideas of strong and weak ties and dense and loose networks (Granovetter 1973; Loury 1977; Coleman 1988; Burt 1992) but these ideas lacked a clear definition about how these terms could be applied to actual social networks of people. Putnam's work has gone on to be influential in many social media studies, including this work, as will be shown in the second half of this chapter.

Social capital theory has been applied by economists in the arena for developing world economics. It has been argued that most of the economic analysis of developing economies focuses upon nation states, corporations and 'rational' individuals. Furthermore it has been argued that there has been little account of civil society and that the ideas embodied in social capital can bridge a gap between people and markets (Woolcock 1998). Woolcock developed the ideas of Putnam in using a definition of social capital which encompasses social protocols between individuals based upon trust and reciprocity (Putnam 1993). Woolcock argued that there was a type of social capital which he called 'linking social capital' defined as the relationships between individuals and groups across different social strata in a hierarchy of status and wealth (Woolcock 2001).

Social capital has also been expressed as an integrative framework for understanding knowledge creation and sharing in organisations (Nahapiet and Ghoshal 1998). It was argued that organisations have advantages for creating knowledge over natural settings because they provide an environment conducive to the development of social capital. Furthermore it was suggested that the creation and development of knowledge is facilitated when; individuals are motivated to participate, there is an underlying social network structure, individuals have the cognitive ability to manipulate knowledge and the ties in the network are positive and strong (Nahapiet and Ghoshal 1998). These assertions, it was argued, amount to an integrative framework making up social capital with the following components; motivations, structural capital, cognitive capital and relational capital. The integrative framework was applied to group level social capital.

Both the World Bank and the Organisation for Economic Cooperation and Development (OECD) have similar definitions of social capital. For these organisations social capital refers to networks with shared norms and values which facilitate cooperation between and

within groups of people (WorldBank 1999; OECD 2001). Specifically for the World Bank social capital is

“the institutions, relationships and norms that shape the quantity and quality of a society’s social interactions” (WorldBank 1999).

These definitions focus on an analysis into the potential social capital has for economic development and societal advancement. However the World Bank also points out that as well as having positive effects, such as increasing productivity, social capital can also have negative aspects if the social group is too narrow and becomes parochial (WorldBank 1999).

It has been argued that much of the discourse on social capital has fixated on the manifestations of social capital rather than the concept itself (Fukuyama 2001). Fukuyama argues that the definition of social capital should focus upon an ‘instantiated informal norm’ which promotes collaboration between individuals (Fukuyama 2001) in a similar vein to Bourdieu (1985). The norms can range from two individuals through to formal religions. However the argument Fukuyama makes is that the norms must exist (or be instantiated) rather than being potential. Furthermore not just any instantiated norms constitute social capital, the norms must lead to cooperation and include virtues like honesty and reciprocity (Fukuyama 2001). Fukuyama’s definition seems at first reading to be pedantic in so far as other theorists are often implying that social capital is instantiated. However this distinction has value for research in social media because there is great deal more potential social capital than is actually experienced by the individual participants in social media. For example in a social media network of around one hundred people the fact that two individuals join the network suggests that they both have equal potential for instantiating social capital. However the fact that there are some pre-existing social bonds between the new participants and existing ones means that the instantiation of social capital may not be the same.

Overall the literature on social capital contains a multiplicity of sometimes conflicting definitions for the term social capital. For example both Lin (2001) and Gargiulo (1999) are broadly in the Bourdieu (1985) tradition with a focus on the individual perspective, rather than the group, and on the underlying social structure as being social capital, rather than social interactions. However many theorists use a similar approach to Putnam (1995, 2000) and base their definitions of social capital on the ideas of social relationships (Adler and Kwon 2002; Cohen and Prusak 2001; Resnick 2001; Tamaschke 2003; Dakhli and DeClercq 2003; Arregle, Hitt and Sirmon 2007; Krishna 2012).

It is plain that there is no undisputed definition of the term social capital, let alone on how to create a measurement. One path to overcoming the inherent difficulties in working with a term overloaded with many meanings can be to adopt a multidimensional approach. One such endeavour was conducted by Scheufele and Shah (2000), who built upon the work of Putnam in conceptualising social capital as referring to networks, norms and trust that give people the means to take collective action (Putnam 1993, 1995). Scheufele and Shah state that there are three dimensions of social capital which are: intrapersonal, interpersonal and behavioural. The intrapersonal dimension refers to a person's life satisfaction. The interpersonal dimension refers to trust in terms of social or generalised trust. The behavioural dimension refers to individual's active participation in civic and political life (Scheufele and Shah 2000). This systematic breakdown used by Scheufele and Shah is again reminiscent of Putnam's bonding and bridging capital analysis. Ideas for examining the different aspects of social capital, such as these, are particularly useful for computer based studies such as the one outlined in this thesis, because it can provide a theoretical framework to analysing different interactions.

In summary the definitions of social capital are quite general and are sometimes treated synonymously with the notion of social interaction. However social capital has the benefit of being a systematised means of understanding the characteristics of social connectedness and has proved to be of use in a range of settings from rural isolation to economic disadvantage. The value that social capital has to this thesis is that it provides a theoretical basis for a measure of social value that can be used to derive metrics for social value in social media.

In the next section the literature on social capital theory will be critically examined.

2.4 Critique of Social Capital Theory

The concept of social capital originated as a sociologic theory and has been exported into everyday language via cognate domains as diverse as psychology and mathematics. Like many terms used in common parlance the original meaning and heuristic value have become diluted. However social capital still retains originality and flexibility because it is a means of assigning value to sociability and human relationships.

There is an important disagreement in the literature on social capital as it relates to the difference between those who view social capital as being an attribute of an individual versus those who consider social capital to be an attribute of a group (Ichiro et al. 2004). Much of the confusion around social capital theory is in regard to its application to different types of problem areas using different units of measurement. Two of the original theoreticians responsible for developing the concept of social capital are Bourdieu (1980) and Coleman (1993), who centred their work on both individuals and small groups. Although there are differences between Bourdieu and Coleman they both focused on social capital as a benefit which can be increased by ties to others. Bourdieu's theories were, in particular, influential in noting that individuals built their relationships for the benefits they might bring in the

future and that social capital is fungible between financial capital, cultural capital and social capital.

The subsequent literature has, for the most part, followed Bourdieu's theoretical guidelines by focusing on the types of resource that can be accrued by means of a social network. In the cognate domain of sociology the idea of social capital has become defined as: a source of control, a source of family enabled benefits and a source of non-family benefits Bourdieu (1980). Coleman concentrated on the use of the term as a source of social control. From this perspective, a range of negative consequence arose, ranging from crime to freeloading (Coleman 1988). After Coleman's contribution a subtle distinction began to emerge as the concept of social capital was exported into other cognate areas, social capital began to be used as an attribute of the community itself. For example, Putnam (1993, 1995) used the idea of 'stock' of social capital possessed by communities and even states. Although Putnam does not directly challenge Coleman's work, the stretching of social capital to cover nation states is not a small step but rather a qualitatively different proposition. Therefore the cause of building a coherent theory of social capital has come from two different perspectives. The first perspective defines social capital primarily as a characteristic of an individual. The individualistic perspective defines social capital as being made of a network of social connections (Bourdieu 1996). The second perspective treats social capital as an attribute of a community, in terms of the networks and relationships which allow individuals to collaborate (Putnam 1993).

The cause of developing a social capital theory has essentially created a schism where there is not even agreement on the central definition of social capital. The schism is due to the fact that the transition of the concept from an individual attribute to a community or national attribute has not been explicitly theorised and therefore has allowed the present state of

confusion about the meaning of the term to arise. In one source, social capital is used to refer to individuals in a family, and in another it refers to a network of traders, and in yet another, as an explanation for the economic decline of cities (Portes 2000). The heuristic value of social capital is diminished as it becomes synonymous with positive aspects of social interactions.

The confusion regarding social capital is evident where two competing definitions are compatible in some instances but not in others. For example, connections allow individuals to gain access to profitable contracts and thereby bypass regulations which are binding to others. In such a situation individual social capital is gained by an ability to undermine community, or collective, social capital. Furthermore, the causes and effects of social capital as a group trait have not been clarified. Bourdieu's tradition of individual social capital has not suffered from this because of theoretical underpinnings given to the concept. In Bourdieu's tradition the sources of social capital were associated with a person's network, while effects were linked to material and informational benefits, clearly separate from the social structures that produced these benefits (MacLanahan and Sandefur 1994; Hagan, MacMillan and Wheaton 1996).

In summary the term social capital has been used to describe a property of both groups and individuals. However the use of the term social capital in association with individuals is the only use clearly theorised by Bourdieu (1985). Fortunately, theoretical problems with the concept of social capital have been somewhat ameliorated by efforts to measure it empirically. For example newspaper reading, trust measures in surveys and participation in community groups have all served as indicators of the effects of social capital (Putnam 1995). In addition the literature in computer based studies of social capital includes studies explicitly examining ties in a social network examined by means of effects such as improved sociability

(Liccardi et al. 2007). However, that is not to suggest that the measurement of social capital is without problems, but merely that as the literature has matured the results of studies have helped to clarify the theoretical concepts.

A range of researchers from diverse disciplines have argued that high levels of social capital in real world communities are positive. For instance, Putnam (2000) has suggested that social capital enables people to collaborate in order to resolve common problems. Social sanctions are also used for enforcing rules where breaches in social norms occur. Putnam also suggested that when people are trusting and maintain interaction, everyday business is easier. He added that social networks can act as a means for the dissemination of information that can assist in the fulfilment of individual and community goals. For instance, people who are well connected in a social network will normally receive useful news first (Putnam 2000).

It has been argued that social capital can negate cultural differences by building a shared identity and new social norms (Daniel, Schwier and McCalla 2003). Furthermore, from the standpoint of organisational management it has been noted that social capital can facilitate improved knowledge sharing due to trust relationships, common values and shared goals (Prusak and Cohen 2001). Social capital can create a range of benefits in diverse type of communities. For example closed communities enable reciprocity and trust to emerge in dense networks where members have frequent interactions in a closed social network (Woolcock 2001, 1998). It has also been argued that communities with high levels of social capital have frequent interactions among members, which can instil the development of social norms through which members are more likely to assist one another, which in turn may improve dissemination of information (Narayan and Pritchett 1997).

Whilst there are undoubtedly potential benefits associated with social capital for individuals and communities, such as better outcomes in education and in business, there are also a

number of potential problems. For instance Portes (1998) outlined a drawback of social capital in terms of the tendency to exclude 'outsiders', and the imposition of restrictions on individual freedom caused by social protocols and collective norms. There are further potential drawbacks evident in situations where group cohesion is achieved by a shared experience of adversity and opposition to mainstream society as is the case with some religious cults and hate groups. Another negative tendency is the so-called downward levelling of social norms which can create a situation where inequalities between members become chronic (Portes 1998).

Some of the issues caused by highly cohesive social networks are that they can exhibit bonding social capital which is not beneficial to the wider society and may in fact create trust among members whilst spreading criminal values and behaviours. Therefore bonding social capital can be manifested in cohesive social networks and not be beneficial to society at large. In some situations, social capital can provide the advantages outlined above but go on to also create a dense network which in turn becomes isolated from the larger social grouping. In situations such as these, the benefit of group membership may not outweigh the disadvantages. For example, the exchange of information in a dense network can be optimal; however if the information itself is restricted the community will not be open to new ideas (Woolcock 2001). In terms of social capital theory it can be argued that these dense networks with high levels of bonding social capital sacrifice the potential benefits of bridging social capital as a means to a wider sphere of information. In addition to the issues associated with types of social capital in groups, it has also been suggested that a high level of bonding social capital can lead to a decline in looser associational ties and therefore a lack of civic involvement (Putnam 2000). Furthermore there is a distinction to be made where connections between individuals in a social network may be aberrant in terms of having a

negative impact on one or both individuals. In social capital theory, typically, there is not a distinction between negative connections in the network and positive ones.

In summary, it is useful to emphasize that social capital with social norms and collective sanctions can be problematic in some contexts.

2.5 Overview of Social Capital Theory

Social capital as a concept is well established with the literature typically using the term to refer to social connections (Adler and Kwon 2002; Cohen and Prusak 2001; Resnick 2001; Tamaschke 2003; Dakhli and DeClercq 2003; Arregle, Hitt and Sirmon 2007; Krishna 2012).

One of the most prominent early theorists was Bourdieu (1985) who focused his analysis on the network structure that enables social connections. The literature contains a number of sources which provide an analysis of the type of connections in a social network (Granovetter 1973; Loury 1977; Coleman 1988; Burt 1992). Arguably the most influential analysis of the connections in a network is that of Putnam (1995, 2000) who posited the ideas of bonding (friends and family) and bridging (acquaintances) social capital. It is Putnam's ideas on bonding and bridging social capital that have particular relevance to this research.

It was also found that there is a contradiction between some of the sources in literature in terms of focusing on social capital as an attribute of an individual only (Bourdieu 1985) or also of a group (Putnam 1995, 2000). The literature contains consideration of the fact that social capital can have negative consequences for an individual if bonding social capital becomes a bar to accessing bridging social capital (Woolcock 2001). However there are several sources that suggest that social capital can be positive in terms of allowing information sharing and collaboration (Daniel, Schwier and McCalla 2003; Prusak and Cohen 2001; Woolcock 2001; Narayan and Pritchett 1997).

The next section contains a discussion of how Social Capital theory has been utilised in the analysis of social media.

2.6 Theories of Social Capital used in Social Media

In this section the literature review is developed to consider how social capital theory has been used in studies concerned with the Internet and social media. The principle research findings in these studies are also examined. This process of considering how social capital theory has been used in different contexts is essential to arriving at an appropriate approach for this research.

The ideas and definitions from social capital theory have been used in a number of computer based studies. For instance the earliest studies included Wellman et al (2001) who made use of the theoretical principles developed by Putnam (2000) in a study examining how the Internet affected social capital. In particular the paper analysed the ideas of network capital, which includes friends, and participatory capital, which includes political involvement. The paper added the idea of community commitment to the theory base and suggested that the Internet could supplement an individual's social capital beyond the usual face to face interactions (Wellman et al. 2001).

In another relatively early study Hampton (2003) published a paper examining tie strength in social media software. The paper used the concepts of aid, information and companionship as indicators of social capital. The theoretical examination of social capital focused on the ideas of networking social capital and bonding social capital laid out by Granovetter (1973) and Putnam (2000) (Hampton 2003),

A study by Rafaeli, Ravid and Soroka (2004) used the idea of cultural capital (Bourdieu 1985) in the context of endeavouring to increase participation in online forums (Rafaeli,

Ravid and Soroka 2004). In addition, the study makes use of Putnam's work which connected social capital to political participation (Putnam 1993, 1995). Rafaeli's contribution builds upon the traditional usage of the term "social capital" by defining "community virtual (online) social capital":

'a collection of features of the social network created as result of virtual (online) community activities that lead to development of common social norms and rules that assist cooperation for mutual benefit' (Rafaeli, Ravid and Soroka 2004).

In addition personal virtual (online) cultural capital was defined as:

'the level to which a person is involved with the virtual (online) community'

One of Rafaeli's contributions to the literature is in drawing a connection between the real and virtual (online) worlds in regard to social capital. Rafaeli's study based on the internet provides a useful early precedent for this study.

In a similar vein to Rafaeli, Ravid and Soroka, Wasko and Faraj (2005) carried out a study into the reasons behind non reciprocal contributions in discussions forums. In this study the definition used was:

"resources embedded in a social structure that are accessed and/or mobilised in purposive action" (Lin 2001)

Wasko and Faraj's study considers why some individuals contribute to a discussion forum whilst others do not. Given the nature of Wasko and Faraj's study it is not surprising that there is emphasis given to sources emphasising the instantiation of social capital. The Wasko and Faraj study also makes use of the idea of social capital residing in the 'fabric of relationships' (Putnam 1995). Furthermore the study examines the hypothesis that social

capital will not translate from the real world to the virtual world (Nahapiet and Ghoshal 1998). It was concluded that social capital does in fact translate to online media (Wasko and Faraj 2005). The Wasko & Faraj study lends further support to the argument that social capital can be analysed in online relationships.

The question about whether or not the internet reduces or increases social capital is a recurring theme in the literature. Best and Krueger (2006) carried out a study which concluded that the level of online interactions positively relates to common indicators of social capital. The paper cites the work of Coleman (1988) as a key source of theoretical concepts such as reciprocity, norms, networks and trust. The measure of social capital used in the study was made up of; trust, reciprocity and integrity. The argument put forward in explanation of the findings suggested that analysis of online social capital should separate out strong tie networks made up of offline friends and family from weak tie networks which were argued as being online networks (Best and Krueger 2006). The Best and Krueger paper was essentially an early exploration of how social capital theory can be related to social media.

In a study examining the effects of internet use on social ties Zhao (2006) examined different types of internet use. The paper's results suggested that there were differential effects between email and chat in respect to social ties. The paper's literature review was focused on Internet use rather than on Social capital theory per se, however the Wellman and Frank (2001) paper is cited. The paper makes use of the terms social ties and connectivity in a manner broadly similar to the discussion regarding social capital.

A study by Williams (2006) took a systematic view of the measurement of social capital in relation to internet use by devising an Internet Social Capital Scale (ISCS) with the intention of measuring 'bridging' and 'bonding' social capital in both the real and online worlds. The contribution of the ISCS to the measurement of social capital is considerable and is discussed

below. The definition of social capital used in the study was based partly upon the idea of social networks with associated norms of reciprocity with the implication that social capital is both the networks and the effect of the network (Putnam 2000). In addition the Williams (2006) study also made use of the idea of social capital as a process with cyclical patterns in communication technology that comprise socio-technical capital (Resnick 2001). The study makes a distinction between the individual and the community level social capital in the Putnam tradition (Williams 2006). This study has particular importance because firstly it signposts the importance of network structure and secondly because it provides an evidence based assertion that technology can add to social capital. The ISCS was therefore used as a measure of social capital against which to validate the proposed metrics in this study.

Liccardi et al (2007) carried out an analysis of student online social networks without specifically setting out, or citing, a definition of social capital. However the study used a number of concepts common to social capital theory such as the notions of ‘relationships’, ‘trust’ as well as ‘strong and weak ties’. The notion of strong and weak ties corresponds to ‘friend’ and ‘friend of a friend’ in social networks such as Facebook. The study essentially argues the importance of social learning underpinned by online relationships (Liccardi et al. 2007).

The idea of analysing and developing online communities is a common theme in the literature. Smith (2008) contributed a paper reporting on a mathematical method of analysis with the goal of maximising participation in social media. The main sources for social capital theory in this paper are Putnam (2000) and Lin (2001). The concepts on bonding and bridging social capital cited from Putnam (2000) were used as well as with the idea of actual and potential social capital Lin (2001). Smith argued that an Implicit Affinity Networks (IAN) and Explicit Social Networks (ESN) are a means to formalising the analysis of social

capital (Smith 2008). The paper is significant to this research because it represent an attempt to formalise the analysis of social capital in social media. However the study does not have a field study to illustrate the application of the model. It therefore remains only a theoretical construct.

In two studies by Ellison, Steinfield and Lampe in (2006) and (2007) into the effects on social capital of Facebook cited several sources (Coleman 1988; Adler and Kwon 2002; Resnick 2001) in arriving at a working definition for the term Social Capital. The Ellison study quoted a definition of social capital as being:

“The sum of the resources, actual or virtual, that accrue to an individual or a group by virtue of possessing a durable network of more or less institutionalised relationships of mutual acquaintance or recognition” (Bourdieu and Wacquant 1992).

The study concluded that there were benefits in terms of increasing social capital from Facebook (Ellison, Steinfield and Lampe 2007). In a follow on study in 2008 Steinfield, Ellison and Lampe examined social capital as it relates to self-esteem. The terms social capital was explained as being generally understood to mean the benefits we receive from our social relationships and once again they used the Bourdieu and Wacquant (1992) definition (Steinfield, Ellison and Lampe 2008).

In a study into the social capital of Facebook users by Valenzuela (2008) the focus was on the relationship between social capital and student participation in University life. The 2008 study and another in 2009 discusses the challenges of arriving at a working definition of the term social capital (Valenzuela 2008; Valenzuela, Park and Kee 2009). The definition which was used for their research was based upon multi-dimensional approach suggesting that

social capital is made up of the dimensions of intrapersonal, interpersonal and behavioural (Scheufele and Shah 2000).

Much of the literature on social capital and social media has focused on internet use in general, or more recently on Facebook in particular. However Pfiel, Arjan and Zaphiris (2009) carried out an interesting study into the use of MySpace from the perspective of difference in age groups. In common with many of the sources in the literature the study made use of Putnam (2000) as a source for theoretical ideas such as bonding and bridging capital as well as Granovetter (1973) for the ideas of strong and weak ties, which Pfiel, Arjan and Zaphiris hybridised. The resultant theory is that strong ties are related to bonding social capital and weak ties to bridging social capital. Using these theoretical ideas the study reported that younger users had larger networks with weaker ties which were fairly homogenous, whilst older users had small networks of strong ties but with a more heterogeneous group of people (Pfeil, Arjan and Zaphiris 2009).

In a study examining online relationship in terms of the strength of ties in social media by Gilbert and Karahalios (2009) there is again no specific definition of the term social capital. However the term is used several times and social capital theorists are cited (Burt 1992; Lin 2001). The research addresses the theory laid out in (Granovetter 1973) regarding the measurement of tie strength (Gilbert and Karahalios 2009) which is a key idea in the process of analysing social connectedness in virtual social networks.

Some of the literature regarding social capital in online interactions considers the idea of whether or not the internet can facilitate the development of networks and hence increase social capital. In a study into levels of internet use and how they relate to real world community participation Wellman et al (2001) uses Putnam's definition of social capital as relating to political participation. However, Wellman et al counter Putnam's suggestion that

social capital is in decline by suggesting that online interactions are being substituted for real world interactions which therefore appear to decline. In a sense, then, the argument outlined in Wellman et al is that Putnam has inadvertently measured a decrease in social capital due to the increase in the use of the internet (Wellman et al. 2001). Wellman et al are, in common with other sources, arguing that online communities are a continuation of real world communities and that, since the advent of the internet, social capital is composed of both virtual and real world connections.

In the last few years the dominance of Facebook in the realm of online communities has been reflected in some of the literature; for example Burke, Kraut and Marlow (2011) examined how Facebook affects social capital. The definition of social capital used in the study was based on the work of Bourdieu on social capital as being made up of actual and potential resources in 'relationships of mutual acquaintance' (Bourdieu 1985). Burke et al also make use of the concepts of 'bonding' and 'bridging' social capital referring to close friends and looser acquaintances (Putnam 2000). Burke, Kraut and Marlow (2011) argue that bridging social capital can be a source of more effective social capital as the ties are weak. The argument is that bonding social capital contains too much redundant information in the tradition of the ideas of Burt and counter to Coleman although these sources are not cited in the paper (Burt 1992). The paper concludes that Facebook use is not associated with Bonding social capital but that bridging social capital is increased by Facebook person to person interaction (Burke, Kraut and Marlow 2011). The source examines ties in some depth, stating amongst other points, that weak ties are more likely to be a source of negative information than strong ones. However there is no clear conclusion in the study regarding the relative merits of strong versus weak ties in terms of social capital.

In more recent studies building on earlier work, Ellison, Steinfield and Lampe (2011; 2014) cite the main sources of social capital theory as being (Bourdieu 1985; Coleman 1988) and in addition (Burt 1992; Putnam 1995; Lin 2001). The study was concerned with the idea of bonding and bridging social capital (Putnam 1995) and also contributed to the debate concerning strong and weak ties based upon (Granovetter 1973; Burt 1992). The definition of social capital used in the source is once again the actual and potential resources in 'relationships of mutual acquaintance' (Bourdieu 1985) used by (Burke, Kraut and Marlow 2011) and (Ellison, Steinfield and Lampe 2011). Therefore Ellison's (2011; 2014) contributions uses a traditional definition of social capital that helps us to arrive at a working definition for the term in the sphere of computing, a definition which must take account of the nature of the ties in the network.

Virtually all of the literature on social capital and social media accepts the software as an entirety and as a causal agent in either promoting or inhibiting social capital. However one study in particular examined the effects of elements of the Facebook interface in an effort to understand the effects of social media design on social capital Yoder, Hill and Stutzman (2011). The theoretical underpinnings for this study included bonding and bridging social capital from Putnam (2000) and reference to the work of Ellison, Steinfield and Lampe (2007) as well as Steinfield, Ellison and Lampe (2008) and Valenzuela (2008). The study suggested, perhaps unsurprisingly, that interaction elements that supported one to one communication were positively associated with social capital (Yoder, Hill and Stutzman 2011).

The tradition of using social capital as a means for analysing people's political participation and civic engagement has been carried on into the sphere of social media. In a study by Gil de Zúñiga (2012) it was found that using social media as a source of information was a

positive predictor of people's social capital and civic participation both on and offline. The source cites (Bourdieu 1985; Coleman 1990; Lin 2001) but focuses on Lin's (2008) definition of social capital as being the embedded resources, that can be accessed, within a social network (Gil de Zúñiga, Jung and Valenzuela 2012). The implicit assumption in Lin's definition i.e. the utility of social capital or the instantiation assertion (Fukuyama 2001) is that it is necessary to focus on the actual rather than the potential. The theoretical point under analysis here is the idea that it is possible to have potential access to social capital but that access might not be instantiated in all instances. Whilst this point may seem to be rather rarefied it has linkage to the notion of cultural social capital (Bourdieu 1985) which, amongst other things, is a measure of potential access to expertise. The distinction between utility (or instantiation) and potential social capital may turn out to have particular relevance to social media where the messaging to individuals can far outweigh their desire to respond.

In another pair of papers by Junco (2012b, 2012a) outlining a Facebook study this time in regard to students in Higher Education it was found that there were both positive and negative aspects of Facebook use in the academic arena. The theoretical basis for Junco's study comes from (Ellison, Steinfield and Lampe 2007, 2011; Valenzuela, Park and Kee 2009) and therefore implicitly uses definitions of social capital derived from the 'sum of resources' theory of (Bourdieu and Wacquant 1992) and the 'intrapersonal, interpersonal and behavioural' multi-dimensional approach of (Scheufele and Shah 2000). The study is primarily concerned with measuring the effect Facebook has on Grade Point Average at university, so it is not surprising that there is no working definition cited in the source. However the Ellison studies into the effect on social capital of Facebook are cited as a rationale for the work (Junco 2012a, p. 187–198, 2012b, p. 162–171).

Antoci, Sabatini and Sodini (2012) put forward a logical framework for assessing the role of the internet in the evolution of social participation, both online and offline. A key source for social capital theory used in the paper is Putnam's (2000) ideas regarding bonding and bridging social capital. The paper outlines the argument that social media can support the strengthening of bonding and bridging capital as well as allowing the crystallisation of weak or latent ties and in support of this assertion cites Ellison, Steinfield and Lampe (2007), Steinfield et al (2009), Gilbert and Karahalios (2009) and Burke, Kraut and Marlow (2011). The study aimed to create an evolutionary framework for examining the dynamics of social interaction and concludes that opportunities for participation will lead to a growing share of the population using social media (Antoci, Sabatini and Sodini 2012), a suggestion that few would refute given the rapid expansion social media use has produced over the last few years.

Hofer and Aubert (2013) carried out an interesting study into how social capital relates to Twitter. The theoretical basis of the paper included the ideas outlined by Coleman (1988) asserting the difference between potential and instantiation. Furthermore in common with many sources in the literature the theory used included Granovetter's (1973) work on network ties, Putnam's (2000) contribution of bonding and bridging social capital and Williams' (2006) Internet Social Capital Scales. The study examined the relationship between followers and followees, and found that number of followers was associated with bonding social capital whilst number of followees was associated with bridging social capital (Hofer and Aubert 2013).

In a study into the dimensions of social capital in Facebook friends, Jung et al (2013) used Williams (2006) ISCS as a means of analysing how online social relationships can facilitate access to resources. The study found that some sub-dimensions of social capital did predict responses to requests for favours from Facebook friends, bonding and bridging social capital

did not (Jung et al. 2013). The theoretical underpinnings for the study came from (Williams 2006) and by extension from (Putnam 2000) and (Resnick 2001, 2004). This study shows an increased sophistication in analysing the constituent parts of social capital rather than significantly adding to the theory base.

A study examining social media in the context of employee work performance using social capital theory as a frame of reference, found that social capital was increased by both social and work related use (Sun and Shang 2014). This study shows an expansion in the literature in the use of social capital theory in social media studies, from the public arena into the organisation arena. The theoretical elements of the Sun and Shang (2014) paper in terms of applying social capital theory to organisations were based upon Adler & Kwon (2002), Tsai & Ghoshal (1998) and Nahapiet & Ghoshal (1998). However in terms of social capital theorists per se, the work of Coleman (1988) and Putnam (1995) were cited. It would seem that the work of Putnam is prevalent in the social media and social capital theory literature, particularly the idea of dividing social capital into bonding and bridging social capital.

Appel et al (2014) conducted an insightful study examining the validity of the ISCS (Williams 2006) and argued the method lacked convergent validity with other measures of social capital such as name generators, resource generators and position generators.

Generators are methods which ask participants to list specific items. In the case of name generators participants are asked to list the names of people that they know, for resource generators the list is made up of resources and position generators relates to the professions of people that they know. However it should be noted that there is no standardised method of generating the lists in social capital theory. In relation to this work name generators would not be an appropriate method because this examination of social capital is specific to software and the software provides a view of the names of the people that participants in a study would

list. In the case of resource and position generators the results would be unlikely to alter in relation to using software in the time period of a study. Furthermore Appel et al (2014) did find a correlation between ‘support through positive interactions’ and William’s ISCS. This correlation, which is relevant to this study, taken together with the lack of a single definition of the term social capital suggests that the ISCS is a valid measure for this study. In summary the Appel et al (2014) approach is not appropriate for a study concerned with software and the lack of single definition of social capital means that the ISCS is valid even under Appel et al (2014) if the goal is to examine interactions.

In the next section we will consider the benefits and problems associated with levels of social capital.

2.7 Critique of Social Capital in Social Media

In terms of the potential benefits of social capital in Social Media the literature suggests that high levels of social capital are beneficial and that Social Media can either supplement or replace traditional access to social capital (Wasko and Faraj 2005; Williams 2006; Ellison, Steinfield and Lampe 2006; Valenzuela, Park and Kee 2009; Gilbert and Karahalios 2009; Wellman et al. 2001; Burke, Kraut and Marlow 2011; Gil de Zúñiga, Jung and Valenzuela 2012; Jung et al. 2013). Furthermore, it has been argued that decreases in social capital in the real world which are thought to be a negative phenomenon (Putnam 2000) have actually been replaced by structures of relationship constructed in online media (Wellman et al. 2001). The idea that computer based communication can provide access to cultural social capital in terms of access to experts is also an example of the potential positive aspects of social capital (Gil de Zúñiga, Jung and Valenzuela 2012). It has also been suggested that a person’s psychological well-being can be improved by increasing online social capital via Facebook (Ellison, Steinfield and Lampe 2007).

On the other hand, the literature regarding online social capital has in the last few years started to examine the extent to which negative aspects of online social capital are extant. For instance, it has been suggested that students using Facebook's status update facility to make negative remarks is likely to decrease the individuals likability within their social network (Steinfeld, Ellison and Lampe 2008). Furthermore it has been argued that high levels of Facebook use for social purpose was related to poor academic performance (Junco 2012b).

Overall there is an assumption in most of the literature regarding social capital in computer based studies that social capital is, by and large, positive and that computer technology can replace, or supplement, traditional access to sources of social capital. However, it is evident that many of the issues traditionally thought to affect social capital in networks in the real world have been carried into the virtual world.

2.8 Analysis of Key Contributors to the Literature

This section contains an analysis of the key contributors to the literature and sets out; a working definition, commonly held key concepts and an explanation of how this thesis relates to the literature.

In order to reflect on the main theorists in the 'social capital and social media' literature, an analysis was carried out. The process involved examining each of the social media papers and counting the social capital theorists cited and then creating a word cloud. The subsequent word cloud shows the most cited reference as the largest 'word'. It is essential to bear in mind that the word cloud does not illustrate the impact of the theorists, however it does provide a visualisation of the instances of referencing which is a useful tool for reflection.

The findings are illustrated in the word cloud shown in figure 2.1 below.



Figure 2.1 Word cloud of social capital theorists referenced in social media studies

The word cloud shows that Putnam (1995, 2000) is the most cited source largely due to the prevalence in the use of the concepts of bonding and bridging social capital in the literature. Therefore Putnam’s concepts of social capital being made up of bonding and bridging social capital was utilised in this research.

Another significant source is Bourdieu who is frequently used to provide a definition of social capital, which is:

“the aggregate of the actual or potential resources which are linked to possession of a durable network of more or less institutionalized relationships of mutual acquaintance or recognition” (Bourdieu 1980, 1985).

Bourdieu’s definition is clear and comprehensive it is also widely referenced in a range of sources including Putnam. The definition makes reference to both actual and potential resources which are of value when considering the actual and potential resources available in

software. Because the Bourdieu definition is relevant to the work of Putnam and to software it was used as the definition of social capital relevant to this thesis.

Another frequently cited source is Granovetter (1973) who made an early contribution by examining the characteristics of ties in a network which described as being either strong or weak. This work was influential to other contributors to the literature such as Loury (1977) and Putnam (1995, 2000).

2.9 Overview of main points from the literature

Overall the contributions to the literature on social capital in social media contain a number of significant findings which are at times contradictory. A key finding concerns the notion that social capital in online communities parallels real world social capital (Rafaeli, Ravid and Soroka 2004; Wasko and Faraj 2005). However there is disagreement over whether or not online communication increases (Wellman et al. 2001; Best and Krueger 2006) or decreases social capital (Putnam 1995, 2000). It has also been suggested that Facebook can increase social capital (Ellison et al. 2014). This assertion concerning Facebook and the growth of social capital was examined in the practical work outlined below.

The literature has shown that social capital theory has been applied to a range of social media platforms. For example Facebook by Ellison (2006; 2014) and Junco (2012a). In addition there have also been two studies one using Twitter (Hofer and Aubert 2013) and the other MySpace (Pfeil, Arjan and Zaphiris 2009).

In terms of applying Putnam's (1995, 2000) concepts of bonding and bridging social capital to social media these were related to a study by Pfeil, Arjan and Zaphiris (2009) into ties in a network. In addition it was found by Burke et al (2011) that it was unclear whether or not strong or weak ties were better at providing social capital. Perhaps most significantly to this

study, which uses Facebook, it was found by Liccardi et al (2007) that Facebook friends relate to ties in a network.

The literature also contains a valid and widely used survey for measuring online social capital known as the ‘Internet Social Capital Scales’ (ISCS) (Williams 2006). The ISCS survey has, for example, been used in the creation of the Facebook Intensity Scale (FIS) by Ellison (2006; 2014) and also by Jung (2013) to measure social capital.

This thesis is concerned with advancing the usefulness of social capital theory in the realm of social media software analysis. There are two ideas that this research adds to the literature on social capital in social media which are;

- To create and validate social capital metrics which are related to the concepts of bonding (strong ties) and bridging (weak ties) social capital.
- To implement and test a social capital metrics in software with an extensible architecture

2.10 Chapter Summary

This chapter contains an examination of the main elements of social capital theory. It is suggested that social capital theory provides a framework for understanding the structure that exists in a social network. Furthermore, social capital theory provides concepts useful for analysis of online social interactions, such as bonding and bridging social capital. In addition the current literature in computer based studies which use social capital theory was reviewed. It was found that there are a number of studies which have advanced the understanding of how computers can support social capital. This chapter also contained an outline of the tendency in the literature of both traditional social capital theorists and those working in the computer based studies to see social capital as a broadly positive concept. It has been argued that there are a number of downsides to social capital which are as evident in the virtual world as in the real world. The chapter concluded with a summary of the key concepts from

social capital used in the social media studies. In the next chapter contains a consideration of how social capital can be measured in social media.

Chapter 3 Measurement of Social Capital in Social Media

3.1 Overview

The literature review in this thesis contains three aspects which are: 'social capital theory', 'social capital in social media' and the 'measurement of social capital in social media'. The literature review is given over two chapters. This chapter presents the third element of the literature review namely the 'measurement of social capital in social media'. Therefore this chapter contains an analysis of the metrics used for social capital. The analysis in this literature review seeks to determine: precedence, omission, guidance and key findings for this research. The chapter is also intended to provide a comprehensive view of scholarship in this area.

3.2 Social Media

The term Social Media, gained popularity from around 2005, and is used to describe the various forms of media content, created by end-users, which are publicly available on the web (Kaplan and Haenlein 2010). In essence, Social Media is a term applied to web-based applications which allow the creation and sharing of User Generated Content (UGC) as opposed to the traditional model of owner of the site publishing content. In the literature the term Social Media is often used synonymously with the terms 'Social Network Site' and 'Social Networks'. However the term 'Social Media' is used in preference to 'Social Network' in this thesis to make a clear distinction between the term 'Social Network' which is often used to imply the social graph underpinning such sites (as well as the software itself). Furthermore it is suggested that the term Social Media more fully describes the range of text, image, audio and video that are commonly posted to these sites. Therefore in this thesis 'Social Media' is used to refer to the software, whilst Social Network is used to refer to the

underlying structure of social interactions and connections (also referred to in this thesis as the 'social graph').

3.3 Measuring Social Capital in Social Media Research

As seen in the previous chapter across the literature, in both real world and virtual world studies, there is no consensus on a definition of social capital let alone how to arrive at an accurate measurement of social capital. Social capital theory enables us to readily intuit the idea of social value for an individual or a group but measuring social capital either qualitatively or quantitatively has proven to be a complex. This complexity in measuring social capital has led to the development of a range of different approaches and metrics which are reviewed in this section in a chronological order.

An early study conducted by Wellman and Frank (2001) measured social capital in relation to the internet. The purpose of the study was to determine how the internet affected social capital. The study made use of the results of a survey of 39,211 visitors to the National Geographic Society Web site, which was one of the first large-scale web based surveys. The results of the study suggested that Internet interactions supplement face to face communication without either increasing or decreasing social capital. However the study also indicated that heavy internet use was associated with increased involvement in politics and voluntary organisations. Paradoxically, perhaps, the study also suggested that heavy internet use was associated with lower levels of commitment to online communities. The study presciently concluded that Internet use was becoming subsumed into everyday life (Wellman et al. 2001).

The Wellman and Frank (2001) paper was the first Internet related examination of social capital. The study provided an outline pathway for subsequent contributions. The scale of the study taken together with the centrality of the importance of social capital is significant.

However the study by Wellman and Frank (2001) study predates the mid 2000s and therefore misses the impact of social media becoming mainstream.

In a study aimed at verifying the effects online communities in a specific geographic area, Hampton (2003) produced a pioneering paper using a mixture of survey and ethnographic data taken from a suburb in Toronto. The findings in the paper suggested that online communities supported increased weaker ties without weakening strong ties. In the study it was suggested that the internet caused an increase in social capital which was measured as aid, information and companionship (Hampton 2003).

The first study of social capital in relation to social media was conducted by Rafaeli, Ravid and Soroka (2004). The study considered social capital at the community level by examining density of ties as well as measuring personal social capital. Rafaeli, Ravid and Soroka (2004) used as a starting point a metric of social capital that attempted to map the actual community activities onto social networks, using network density (ties in the network), boundedness (how closed the community is), range (how wide is the range of relationships), and strength of ties (how wide and strong are connections between people) (Wellman 1997). Rafaeli, Ravid and Soroka extended the framework put forward by Wellman by including so called 'lurkers' in the study. In other words the Wellman study is concerned with posting and commenting messages whilst Rafaeli extends this approach to include group members who have simply read the messages. In the study the metric was standardised by assigning different values to different types of interactions and in particular assigning differential values to reading posts.

There were two formulae put forward by Rafaeli, Ravid and Soroka (2004) for calculating community social capital and personal social capital. Firstly, community social capital can be calculated as shown in Formula 3.1.

$$\text{Density}(\text{group}) = \frac{2 * L}{N * (N - 1) * P}$$

Formula 3.1 Group Density (Rafaeli, Ravid and Soroka 2004)

Let P be the total number of postings in a given community, N be the number of users and L the number of dyadic links.

Secondly personal social capital was measured in the study as show in Formula 3.2.

$$\text{Density}(\text{individual}) = \frac{LU}{(N - 1) * P}$$

Formula 3.2 Individual Density (Rafaeli, Ravid and Soroka 2004)

Where LU is the number of actual links, N is the number of users and P the number of possible links.

The metric devised by Rafaeli, Ravid and Soroka (2004) was applied to discussion forums in a University and used to measure strategies for increasing participation in non-posting (or lurking) students. The results reported in the study showed a significant correlation between social capital and participation (Rafaeli, Ravid and Soroka 2004).

The Rafaeli, Ravid and Soroka contribution to the literature is valuable for a number of reasons. For instance the study extends some aspects of social capital theory in the tradition of regarding social capital from both the individual and group points of view (Putnam 2000). This contribution proffers an extension to the definition that is applicable to the virtual world. Furthermore, the study utilises computer data which has the advantage of capturing actual interactions rather than reported interactions via surveys. However it is suggested that there is a weakness in the paper, which is common in the literature, namely that it does not contain a compelling argument for linking ideas such as the proportion of user links data to the concept of ‘personal social capital’. There is a similar shortcoming in the analysis of group’s

total interactions as being a proxy for group social capital. Overall Rafaeli, Ravid and Soroka (2004) provide precedence to this work for the use of social media data in measuring social capital. However the measurement of social capital in this study was not validated against another reliable measure.

There has been another study along similar lines to Rafaeli, Ravid and Soroka in so far as it also examined a discussion forum, but instead focused on both social capital and knowledge contribution in a network of practice (Wasko and Faraj 2005). This study used the integrative framework put forward by Nahapiet and Ghoshal (1998) made up of: individual motivations, structural capital, cognitive capital, relational capital. Wasko and Faraj (2005) extended the Nahapiet and Ghoshal study by applying the theory to the individual level, from the original group setting, and therefore adapted the network density and centralisation to the individual level. It was argued that the individual's place in the network influences their propensity to contribute to the network. The method employed in the study was to examine a discussion forum used by legal professionals for a four month period and to determine an individual's degree centrality on the network. In the second phase of the study a selected sample of contributors were surveyed. The degree centrality score was measured using UCINET 6 (Borgatti, Everett and Freeman 1999). The survey was constructed using the integrative framework put forward by Nahapiet and Ghoshal (1998) and included a measure of perceived helpfulness of contribution to the network (Wasko and Faraj 2005).

In the Wasko and Faraj (2005) study it was reported that there was no significant relationship between some components of social capital such as cognitive capital and relational capital and helpfulness of contribution in the discussion forum. However it was found that there was significant relationship between structural capital (network centrality) and helpfulness of contribution (Wasko and Faraj 2005). This study represents a significant contribution to the

literature for a number of reasons. Firstly the study attempts to apply systematic analysis of social capital, based upon multiple factors which were originally designed for a group setting to individuals. The study also uses a combination of interaction data gathered from the discussion forums and survey data. Perhaps most significantly for this thesis, the results indicate that there can be a relationship between social network structural characteristics and perceived utility amongst the users of the discussion forum. However the study also indicates that some of the key aspects of social capital that were used had no bearing whatsoever. The logic of this finding suggests that social capital, as defined in the study, has no relationship to helpfulness of contribution as perceived by the user. Overall it is suggested that this leaves two open issues. The first issue is concerning the usefulness of the social capital metric used. The second issue concerns whether or not user perceptions of the content of contributions is a reliable measure likely to be generalisable to other social media contexts. However the approach of using two sources of data, i.e. user interactions and surveys, to analyse social capital provided useful guidance for this study.

There have been a number of survey based approaches to measuring social capital. Arguably the most influential approach has been the Internet Social Capital Scales (Williams 2006). The Williams (2006) paper was based on the theoretical work of Putnam (2000) and Resnick (2001), and as such utilises the ideas of ‘bonding’ and ‘bridging’ social capital. Williams argues that the presence of a social network is important as a causal mechanism in the formation of social capital (Norris 2002). The Williams (2006) study outlines a survey comprised of twenty statements (see Appendix A). The study also included a validation of the survey based upon 884 participants as well as reliability. Overall the contribution made by Williams (2006) is significant as it is theory based and consistent with other criteria related to the concepts under examination in this thesis. Therefore the study provides a strong argument that the resultant measures are related to social capital. The ISCSs has been

influential in other subsequent studies as will be outlined below. Therefore Williams (2006) ISCS were used in this research to validate the proposed new metrics.

Some sources have argued that any harmful effects of the Internet on social capital are restricted to bonding social capital, and that it is possible to construct new ties with people with whom one has never met. The argument is that the Internet may allow people to generate new bridging social capital. For example, Best and Krueger (2006) found that the amount of time users spent online constructing new ties was a predictor of trust, reciprocity and integrity which, it was argued, are dimensions of social capital. Furthermore it has been argued by Zhao (2006) that the effect of Internet use on social capital is dependent on how the Internet is used. For example, Internet use which is dependent on interactivity with other users has a positive effect on social capital whilst uses such as video streaming do not. This finding suggests, what is perhaps a fairly obvious, i.e. communication is better than consuming information for social capital. However by closely examining the logic the finding supports the idea that social capital is being measured rather than just the value of using the actual software.

The Zhao (2006) study made use of a survey and measured social capital in terms of number of friends. However it will be argued in this thesis that number of friends in social media is not a valid proxy for social capital. Findings described below in the practical work of this research, which used the new metrics, suggested that this is too simplistic a measure.

Another important contribution to the literature was made by Ellison Steinfield and Lampe (2006) who conducted a study into the effect on social capital of Facebook. The study made use of ideas developed by Coleman (1988) that use a definition of the term social capital as being the resources created via relationships amongst people, and the idea that social capital can be both 'actual' or 'virtual' (Bourdieu and Wacquant 1992). The study set out to

investigate whether student use of Facebook helped or hindered in the formation and maintenance of social capital. Furthermore one of the research questions specifically set out to examine the relationship between Facebook use and social capital (Ellison, Steinfield and Lampe 2006).

The methodology employed involved conducting a survey on 286 students. Crucially the software usage statistics were derived from self-reporting of participants in the survey.

Facebook usage was used as a metric of social capital with survey statements such as 'I feel I am part of the MSU community', where MSU referred to a University. The study adapted Williams (2006) ISCS criteria for Facebook analysis. The survey in the study was tested for reliability but not for correlation with Williams (2006) ISCS statements. The revised statements were examined for validity of use for Facebook, by means of interviews with volunteers, and named the Facebook Intensity Scale (Ellison, Steinfield and Lampe 2006).

The results of the Ellison, Steinfield and Lampe (2006) study included the finding that people with high levels of Facebook use reported significantly higher levels of bridging social capital than those with lower levels of Facebook use. In addition bonding social capital also significantly predicted the intensity of Facebook use. Furthermore it was found that general internet use was not a significant predictor of either bonding or bridging social capital (Ellison, Steinfield and Lampe 2006). It is worth emphasising that these correlations are between Facebook use and social capital, unlike Zhao (2006) who examined number of Facebook friends as social capital.

In a paper following on from the 2006 paper Ellison, Steinfield and Lampe (2007) re-examined the 2006 survey data to explore social capital in relation to how a user could stay connected to a community (a characteristic referred to in the study as maintained social capital). The additional analysis suggested that Facebook use enabled users to maintain

social ties with people that they knew prior to articulating to University. Therefore across the two papers it was suggested that social capital has three elements i.e. 'bonding', 'bridging' and 'maintained' capital, and that it appears that Facebook use is associated with increases in all three types of social capital. Furthermore it was argued that Facebook might enable users to convert latent ties into weak ties in so far as the software allows students to view other users who might be useful to them (Ellison, Steinfield and Lampe 2006, 2007). The Facebook software does this by allowing users to search the publicly viewable database of user profiles and also by prompting users with 'Friends Of A Friend' (FOAF) tips. Overall the early work of Ellison, Steinfield and Lampe (2006 & 2007) and there is a later contribution in 2011, provides guidance and precedence for this thesis in terms of examining Facebook using social capital theory.

In a study conducted in order to examine the effects of Social Media on the social capital of college students Valenzuela (2008) also reported in (2009), used a web survey of 2,603 college students. The study found a moderate positive relationship between Facebook use and student's social capital as measure by life satisfaction, social trust, civic participation and political participation. The paper concludes that there is a link between online social networking and social capital (Valenzuela 2008).

In another survey-based assessment of social capital, Steinfield, Ellison and Lampe (2008) reported on the results from two surveys undertaken a year apart at a large American University. The results of the survey, it was argued, indicated that social media web sites such as Facebook were particularly useful in enabling students to maintain ties whilst transitioning from school or work to University. The study concluded that intensity of Facebook use in the first year strongly predicted bridging social capital in the second year of study (Steinfield, Ellison and Lampe 2008). However the results from these two interesting

studies share a limitation in so far as they rely on self-reporting in a survey. Whilst survey methods are undoubtedly useful for studies into social capital, there are likely to be shortcomings in user's recollections and perceptions as opposed to data from actual interactions.

Smith (2008) published a paper containing a mathematical model for examining social capital in social media web sites. The model involved evaluating nodes in a social graph for their relationships and attributes. The intention of Smith's study was to create a quantitative model that would assist decision support on how to maximise participation in a social network. The study, therefore, has a similar goal in mind to that pursued by Rafaeli, Ravid and Soroka (2004) and Wasko and Faraj (2005). The model represents an attempt to formalise online social capital and encompasses the ideas of bonding and bridging social capital outlined by Putnam (2000). Smith used the idea of Implicit Affinity Networks (IAN) based upon similar attributes, such as shared hobbies, between nodes without a specific link, whilst the idea of Explicit Social Network (ESN) is based upon an actual link between the individuals.

Smith (2008) suggests bonding social capital can be defined between nodes i and j as the product of strength of the IAN (s_{ij}^{IAN}) edge by the strength of the ESN (s_{ij}^{ESN}) edge, i.e.

$$\text{bonding social capital } (i, j) = s_{ij}^{IAN} * s_{ij}^{ESN}$$

Formula 3.3 Bonding Social Capital (Smith 2008)

Furthermore bridging social capital between two nodes i and j is simply $1-s_{ij}^{IAN}$ multiplied by (s_{ij}^{ESN}). Smith argues that the more dissimilar the two nodes are, the larger the potential for bridging social capital. Therefore actual bridging social capital between two nodes can be defined as the product of the bridging social capital of the 1- IAN edge by the ESN edge as shown in formula 3.4.

$$\text{bridging social capital (i,j)} = (1 - s_{ij}^{\text{IAN}}) * s_{ij}^{\text{ESN}}$$

Formula 3.4 Bridging social capital (Smith 2008)

Smith states that if IAN and ESN tie strengths are 0 then there is no bridging social capital. On the other hand if both IAN and ESN are 1 there is no bridging social capital as the individual are homogenous (i.e. no significant dissimilarities in knowledge). Bridging social capital is at a maximum of 1 only when ESN is 1 but IAN is 0 (Smith 2008).

The notion of homogeneity being a negative factor goes to the heart of Smith's exposition and can be explained in the situation where an individual is seeking employment. Smith's usage of the term 'homogeneity' suggest that if you are engaged in a job search you will likely already know what those closest to you know about the job market. Therefore, the potentially more useful advice may come from individuals that you are in not in a homogenous grouping with.

Overall Smith's mathematical model makes an insightful contribution to the literature showing a method of applying the theoretical concepts in social capital to social media via a mathematical model. However as Smith's paper merely outlines a model and does not include results it is difficult to fully evaluate this approach. Furthermore there is no indication in the paper of how the Social Capital metric was to be validated. In other words there is no data to suggest the measure is actually social capital and not some other characteristic of a social graph. In terms of this study, Smith (2008) provides a valuable precedent for formalising social capital theory in a mathematical formula.

In a paper titled 'Bowling Online', referencing Putnam's (2000) influential article titled 'Bowling alone', Steinfield et al (2009) examined a company internal (Intranet) social media platform known as 'Beehive'. The study involved adapting the survey using the Facebook

Intensity Scale (Ellison, Steinfield and Lampe 2006, 2007) which in turn was adapted from the Internet Social Capital Scale (Williams 2006). The study concluded that intensity of use of the Beehive software was associated with increased bonding social capital presumably due to enabling access to new people and to expertise (Steinfield et al. 2009). The Steinfield et al (2009) paper makes a significant contribution to the literature in terms of adapting the methodology for use in Intranet social media platforms. In terms of this research the key finding that social media can support an increase in bonding social capital over time was examined in the practical work.

Whilst many studies have made use of Facebook as a case study, Pfeil, Arjan and Zaphiris (2009) used MySpace. The MySpace social media platform reveals the structure of the data in the HTML of the user web pages whereas Facebook does not. This means that it is relatively straightforward to create a script that makes an HTTP request to the MySpace site and can then extract the data from the returned result. As opposed to Facebook pages which contain only display information which means that in order to extract data from the site the API must be used.

The Pfeil methodology involved creating a web script used in order to interrogate the data from MySpace for 70 teenagers and 70 older users. The study involved carrying out a content analysis of the data as well as extracting age data and number of friends. The metric used for social capital was 'number of friends' and it was concluded that teenagers tended to have more friends. It was also suggested that older users might increase their social capital by revealing more personal and emotional information (Pfeil, Arjan and Zaphiris 2009).

The Pfeil, Arjan and Zaphiris (2009) study addresses the tendency in social media studies to analyse Higher Education students' behaviour and makes a welcome contribution by considering demographic dimensions to online social capital. However, as the results were

not validated against any other measure, it is by no means clear that number of friends listed on a social media site is an accurate measure of social capital. The approach of using number of friends as a proxy for social capital was also used by Zhao (2006), described above. However, to reiterate, the findings outlined below in the practical work suggest that this is not a valid approach.

In a study combining network analysis with survey data, Gilbert & Karaholios (2009) set out to create a predictive model that would map social media data to tie strength. It is worth noting that Gilbert and Karaholios did not suggest that they were engaged in an effort to calculate social capital per se, rather they were determining if it was possible to assess the ties between two individuals based upon social media data. The methodology employed involved using an online survey which required the thirty-five student participants to rate their Facebook friends based on five questions such as ‘how strong is your relationship with this person’. The questions in the Gilbert & Karaholios (2009) study are somewhat reminiscent of the questions in the Williams (2006) Internet Social Capital Scales (ISCS) although the Williams study was not referenced. In addition to the scoring by users of their friends, the testing software also saved a number of data points such as ‘Days since first communication’, ‘Users number of friends’ and ‘groups in common’. It was concluded that social media data can be used to predict tie strength and the authors suggested that these results might be useful for determining influence in social media for use in business and politics (Gilbert and Karahalios 2009).

The Gilbert & Karaholios (2009) study represents a significant contribution to the literature as it employs a novel methodology based upon using actual tie related data and the perceptions of participants. In terms of this research, the dual data approach is important because it enables a cross validation between perceptions and actual data. Furthermore, the

study, in essence, arrives at a metric which, it is suggested, serves as proxy for tie strength in a graph. However it is argued that the Gilbert & Karaholios(2009) approach measures a characteristic of social capital namely tie strength rather than social capital.

In a study also employing a dual methodology made up of survey data and the analysis of server logs, Burke, Kraut and Marlow (2011) examined the social capital of Facebook users. The study examined the actual uses the users made of the software and how that was related to social capital. The study involved collecting data from 415 volunteers recruited by means of a Facebook advertisement. The metric used for bridging and bonding social capital was based upon a five point Likert scale using questions from the Internet Social Capital Scale (Williams 2006). To determine how Facebook use affects social capital Burke, Kraut and Marlow (2011) added the ISCS scores to time on the site. The results indicated that receiving messages from friends was associated with increased bridging social capital but that other uses were not. In particular sending messages to friends was not associated with increased social capital (Burke, Kraut and Marlow 2011).

The Burke, Kraut and Marlow (2011) study makes a useful contribution to the literature by examining the use of the software and how that relates to social capital. It seems an obvious but crucial point that it is possible to use software either effectively or ineffectively and will affect the usefulness experienced by the users. For the purposes of this research Burke, Kraut and Marlow (2011) provides a precedent on using the ISCS with Facebook. However it should be noted that the Burke, Kraut and Marlow (2011) study is not attempting to validate a proxy score for social capital, rather they are attempting to relate the ISCS to use of a social media context.

In another study conducted by Ellison, Steinfield and Lampe (2011), adding to the contribution made earlier (Ellison, Steinfield and Lampe 2006, 2007), a paper was published

suggesting that social information seeking behaviours contribute to user perceptions of social capital. The paper utilised a survey of 450 volunteers who were under-graduate students. The metric for bridging and bonding social capital were based upon the Internet Social Capital Scales (ISCS) (Williams 2006). One of the points of focus of the study was to examine the co-relation between number of Facebook friends and social capital. The paper asserted that higher numbers of friends was predictive of social capital up to a point. However the effect diminished as the number of friends that subjects had grew larger.

The Ellison, Steinfield and Lampe (2011) study makes a valuable contribution to the literature in terms of the insights into proxy measures of social capital such as number of friends. The findings cast doubt on the approach to measuring social capital via ‘number of friends’ as reported by Zhao (2006) and Pfeil, Arjan and Zaphiris (2009). The practical work in this research outlined below sheds further light on this issue.

In a study similar in nature to Ellison, Steinfield and Lampe (2011) and aimed at analysing the relationship between the Facebook interface and social capital, Yoder, Hill and Stutzman (2011) found that person to person contact was associated with perceived social capital in Facebook users. The Yoder, Hill and Stutzman (2011) study was made up of a survey of 574 students and used the Facebook Intensity Scale (Ellison, Steinfield and Lampe 2007; Steinfield et al. 2009). The study examined the interface elements in Facebook and found that wall posting (a method of leaving messages for users) was associated with bridging social capital. The metric used for bridging social capital was derived from Ellison, Steinfield and Lampe (2007) which in turn was based on Williams (2006)(Yoder, Hill and Stutzman 2011).

The Yoder, Hill and Stutzman (2011) study makes a useful contribution to the literature as it endeavours to analyse the design patterns in the Facebook software which provides a means

of considering which elements of functionality are most useful in the context of bridging social capital. However there is no evidence of a causal relationship in the findings and it can be argued that people with higher levels of bridging social capital have more opportunities for wall posting interactions than those who do not.

In a study concerned with the creation of social capital and fostering political involvement conducted by Gil de Zúñiga, Jung and Valenzuela (2012), social capital was measured by six items which were; 'feel intimate in the community', 'share community values', 'talk about community problems', 'feel connected', 'help resolve problems' and 'watch out for community members'. The survey was conducted using a 10 point Likert scale and the questions were based on Lin (2008). The results of the study suggested that seeking information via social media was a positive and a significant predictor of social capital (Gil de Zúñiga, Jung and Valenzuela 2012). The Gil de Zúñiga, Jung and Valenzuela (2012) study adds to the Ellison, Steinfield and Lampe (2006) study, suggesting that there is a connection between higher levels of social capital and higher levels of the use of social media. In terms of this research both of these studies provide a precedent for examining social media using social capital theory.

In a paper written by Antoci, Sabatini and Sodini (2012) the authors set out to assess the role of the Internet in the evolution of social participation in the context of the supposed decline of social participation in the previous fifty years in the USA. In essence the source suggests that social media can mitigate the multi-generational downward trend in social capital. The argument in the paper that suggests that there is a long term decline in social capital is based on contributions to the literature by Putnam (1995, 2000) which have been countered by Resnick (2001) and Steinfield et al (2009).

The Antoci, Sabatini and Sodini (2012) study puts forward an analytical framework which focuses on measuring Internet social capital as measured by number of ties as they evolved through time.

In other words increases in the ties in a network will facilitate increases in social capital. Furthermore, since online ties requiring fostering just as traditional ties do, the authors introduced the notion of social capital depreciation (Antoci, Sabatini and Sodini 2012).

The Antoci, Sabatini and Sodini (2012) paper is important as it seeks to construct a method for measuring how the amount of time a person has affects their online social capital.

Although the paper is logically sound it does not present findings based upon utilising the model on an actual data set. Rather the paper intends to provide a model for understanding the circumstances that might lead to an increase to social capital for social media users. This is essentially the same approach as Smith (2008) and shares the limitation of not having supporting data.

In a study based upon surveying Twitter users by Hofer and Aubert (2013) argued that perceived bonding social capital is associated with number of followers whereas bridging social capital is associated with number of followees, but only up to a point. In terms of Twitter use the notion of followers is used to describe the other users who have assented to receive updates from another user's messages and thus provides an audience, whilst the term 'followees' is used to describe the user's connection to other users accounts. In other words, the study found a connection between bonding social capital and the number of recipients of a user's messages, and between bridging social capital and the number of senders to a user's account. The measurement of social capital was an updated version of the ISCS (Williams 2006) adapted for Twitter users (Hofer and Aubert 2013).

In an innovative study by Jung et al (2013) the authors examined the ISCS in sub-divisions beyond bonding and bridging social capital. The methodology employed was an experiment where participants were required to ask a small favour from up to ten of their Facebook friends. The favour being asked was to complete a survey. The findings of the study were that measures of bonding and bridging social capital did not predict responses to favours. However the study examined the ISCS scores in further sub-divisions and found that individual benefit, which is a sub-division of bonding social capital was significantly positively related to favour responses. It was also found that there was a negative relationship between meeting new people, a sub-division of bridging social capital, and responses to favour requests. Furthermore it was found that number of Facebook friends was not significantly related to the number of responses received (Jung et al. 2013).

The Jung et al (2013) study makes a significant contribution to the literature as it employs an innovative approach to understanding the nature of online social capital. Whilst the use of purely survey based methods of research in this area provides a perfectly legitimate way of investigating social capital it suffers from the limitations of self reporting. In essence participants can provide incorrect response for a number of reasons. When asked questions about friends, participants may inaccurately report their relationships in order to cover up social embarrassment. However the favour in the study, which was a request to fill out an online survey, may have altered the participants' evaluation of whether or not it was an important enough favour to bother carrying out. Overall Jung et al (2013) provides an innovative way for going beyond self-reporting in surveys and considers what users are actually able to do with their social capital.

In a study into the intra-organisational use of social media Sun & Shang (2014) found that social (as opposed to work) related use of social media fosters work related use of social

media by enhancing social capital. The methodology adopted in the study was a survey based upon Nahapiet and Ghoshal's (1998). The survey also made use of Putnam's ideas of bonding and bridging social capital (Putnam 1995, 2000) (Sun and Shang 2014). In terms of relevance to this research this study demonstrates the potential future usefulness for organisation of being able to measure social capital.

In another recent contribution to the literature Ellison et al (2014) examined the relationship between bridging social capital and Facebook use for relationship maintenance. The methodology made use of a survey adapted from the ISCS (Williams 2006). The study found that social capital was not correlated with data such as friends but rather by interactions such as posting (Ellison et al. 2014). In term of relevance to this research the Ellison (2014) study provides further evidence that relationships such as friends are not as important as actual interactions. In addition the methodology provides another precedence for making use of the ISCS (Williams 2006) as a means to validate metrics for social capital. The findings in the Ellison et al (2014) study are discussed the practical chapter below which contains a description of how the validated metrics were used.

Given that the aim of this research is to provide metrics that can be implemented in software, particular attention is given to Social Network Analysis (SNA) in the next two sections.

3.4 Social Network Analysis Measures in Social Media

The previous section contained a literature review concerned with measuring social capital in social media. Several of the studies which were reviewed demonstrated that data in social media could be analysed for social capital using SNA. It is argued that SNA has the potential to provide a means for creating automated metrics. Therefore this section contains an explanation of SNA, in general terms rather than specifically for social capital. Thereafter, in

the next section, the discussion will progress on to the theoretical underpinnings for SNA in social media for social capital in particular.

One of the first studies in SNA were carried out by Moreno (1934) on sociometrics and then later by Heider (1946) on triad equilibrium analysis. These ideas were related to graph theory created by Konig in 1936 intended as a formal tool for the study of social structures (Martino and Spoto 2006). SNA involves analysing a graph and its component sub-graphs with the purpose of examining connections between individuals and groups. Graphs are sometimes referred to as sociograms in the context of relationships between people (Moreno 1946). SNA provides a set of descriptive procedures to determine how the graph behaves along with methods to test the appropriateness of experimental propositions (Wasserman and Faust 1994). In order to understand graphs the structure of 'nodes' and 'edges' in the graph corresponding to entities and relationships respectively are analysed. Features of the graph such as degree centrality (closeness to centre by connections) density (the number of connections in the graph) and the idea of strong and weak ties are all concepts frequently used in SNA studies of social media (Valente 2010). Therefore SNA methods have the potential of providing an insight into the structure of the social graph constructed by social media.

Amongst the most commonly used metrics in SNA are degree centrality, betweenness centrality, closeness centrality, clustering coefficient and eigenvector centrality which are explained below.

3.4.1 Degree Centrality

Measuring the network location of a node is known as centrality whilst the number of nodes linked to a given node is known as the degree of the node. Figure 3.1 displays a ‘kite’ network showing a simple graph.

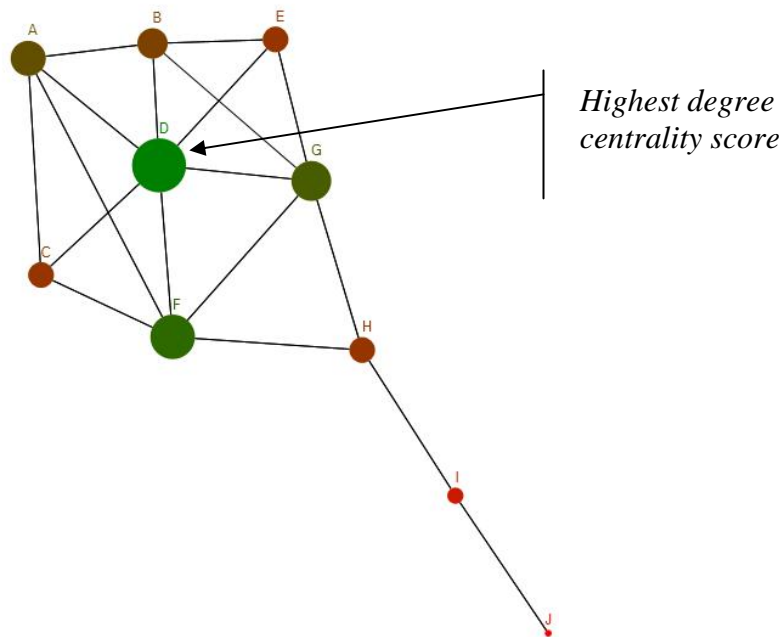


Figure 3.1 Basic graph with node size given by degree centrality

Figure 3.1 illustrates a basic graph with nodes A-J with a variety of edges illustrating relationships between nodes. The number of connected nodes refers to the degree: thus node A has a degree of four. The size of each node is relative to the degree of the node, and so node D is the largest and node J the smallest. The basic graph in Figure 3.1 is useful for outlining some key concepts in SNA. Directedness in a graph indicates a two way connection for example, node I is related to H and J whilst J is only related to I. Degree centrality is expressed mathematically, where v is the vertices (or nodes), as;

$$C_D(v) = \text{deg}(v)$$

Formula 3.5 Degree Centrality (Craven 2013)

Or conceptually as;

Degree centrality of node i = sum of all edges of nodes connected to i

For example node H is connected to nodes G,F and I and thus $C_D(H) = 3$

3.4.2 Betweenness Centrality

Often times when network analysis is applied to social media interactions high degree is considered to be a measure of connectedness, however it is worth noting that whilst node D has the highest degree it is only connected to nodes that are connected to one another.

Betweenness centrality in a graph refers to the number of times a node falls on the shortest path in essence the effect on the graph if the node is removed. Therefore, although node D has the highest degree, node H is the only connector for J and I therefore the betweenness score for H is higher. Betweenness centrality is expressed mathematically as:

$$C_B(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

Formula 3.6 Betweenness Centrality(Craven 2013)

Or conceptually as;

Betweenness centrality of node i = for all relevant nodes; total shortest paths, fraction of shortest paths I , sum over all pairs

For example for node B;

$$C_B(B) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(B)}{\sigma_{st}} = (A \text{ and } E) \frac{1}{2}$$

To summarise, for each pair of nodes, first calculate the shortest path between the nodes. Then for each pair of nodes, determine the fraction of shortest paths that pass through the node in question (here, node v) then sum this fraction over all pairs of nodes. Finally σ_{st} is the total number of shortest paths from node s to node t and $\sigma_{st}(v)$ is the number of those paths that pass through v (Brandes 2001).

Figure 3.2 shows the same type of graph as figure 3.1, but with node size denoting betweenness centrality score.

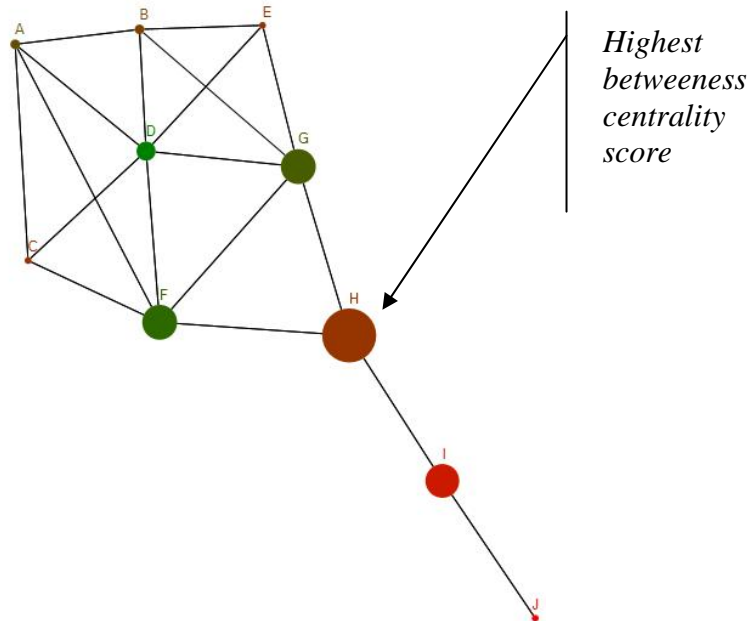


Figure 3.2 Basic graph with node size given by betweenness score

3.4.3 Closeness Centrality

Another key concept in SNA metrics is that of closeness centrality. The concept of closeness centrality is based upon the inverse of idea of farness i.e. the sum of the distance to all other

nodes on the graph. The more central a node is the lower the total distance to all other nodes. Therefore F and G as shown in figure 3.3 will have the highest closeness centrality score as they are close to many nodes (Smith 2003). The closeness centrality of node n_i can be expressed mathematically as:

$$C_C(n_i) = \left[\sum_{j=1}^g d(n_i, n_j) \right]^{-1}$$

Formula 3.7 Closeness Centrality(Craven 2013)

Where g is the number of nodes in the graph, and $d(n_i, n_j)$ is the shortest distance between node n_i and n_j . For example for node I;

$$C_C(I) = \left[\sum_{j=1}^{10} d(I, n_j) \right]^{-1} = \left[\begin{array}{c} 4 + 3 + 3 \\ +3 + 3 + 2 + \\ 2 + 1 + 0 + 1 \end{array} \right]^{-1} = 1/22$$

A network structure that can be visualised as show in figure 3.3

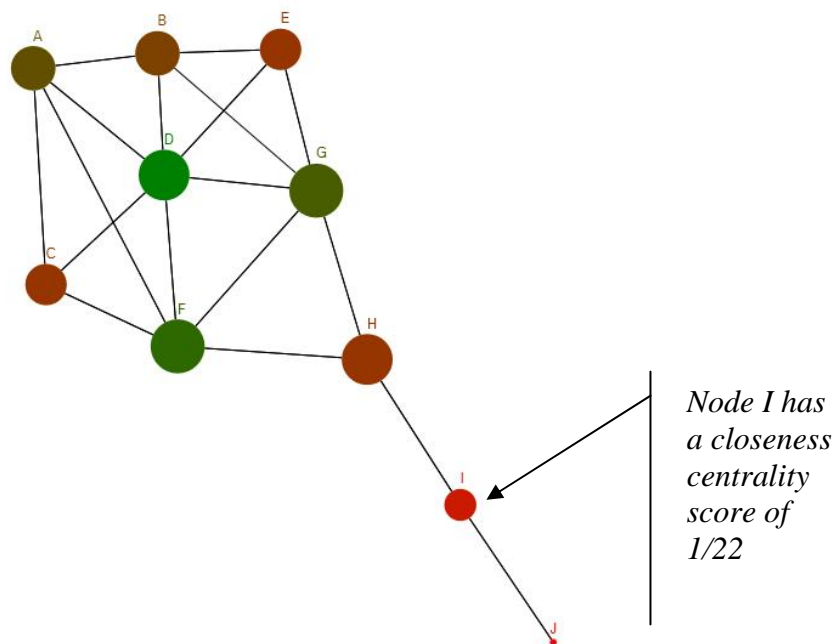


Figure 3.3 Basic graph with node size given by closeness score

3.4.4 Clustering Coefficients

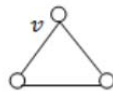
The SNA measure of clustering coefficients is based upon the ratio of number of actual links over number of possible links between neighbouring nodes. This measure is essentially a simple clique score which is expressed mathematically as:

$$CC_v = \frac{\lambda_G(v)}{\tau_G(v)}$$

Formula 3.8 Clustering Coefficient (Craven 2013)

Where $\lambda_G(v)$ is the number of subgraphs of G with 3 edges and 3 nodes, one of which is v , whilst $\tau_G(v)$ is the number of subgraphs with two edges and 3 nodes, one of which is v .

Or visually as;



clustering coefficient of node $v = \text{-----}$



For example for node H;

$$CC_i = \frac{\lambda_G(H)}{\tau_G(H)} = \frac{1}{7}$$

There are a number of variations on clustering coefficient most notably from Opsahl & Panzarasa (2009) who proposed refining the measure using weightings based upon the connections to each of the three nodes. In other words each of the nodes was not treated equally, they were given a start score based on their connectivity to other nodes.

A network structure for clustering coefficient that can be visualised as show in figure 3.4

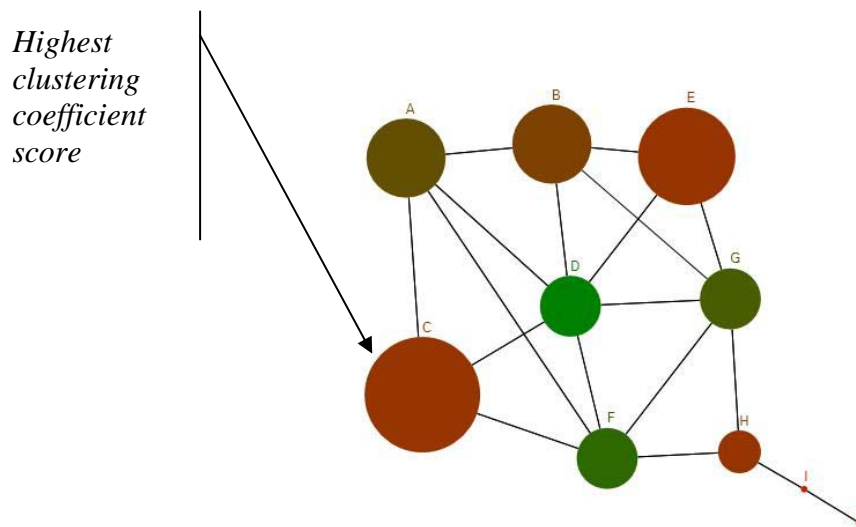


Figure 3.4 Basic graph with node size given by clustering coefficient

3.4.5 Eigenvector Centrality

The concept of eigenvector centrality is essentially concerned with the allocation of popularity scores based upon scores from other nodes. There are a variety of means of calculating an eigenvector centrality such as assigning degree scores to each node and then calculating a new score based upon sharing out the scores from the adjacent nodes. In the case of the kite network node A will start with a given degree score in a graph, then the nodes connected to it will share that degree score, somewhat like allocating pieces of pie. Google's Pagerank algorithm is an example of an eigenvector centrality score (Austin 2013). The logical assumption is that all nodes must first have a score such as degree centrality.

It is worth noting that whenever one node is assigned an eigenvector centrality score based on adjacent scores the new node score affects all other adjacent scores. Therefore the process of calculating eigenvector centrality scores for a graph usually undergoes a pre-determined

number of iterations. Furthermore the eigenvector centrality scores usually increase exponentially and therefore some formulae, for example Pagerank, use a dampening factor.

Eigenvector centrality scores can have a number of formulae that are computed using an adjacency matrix (which is a means of assigning adjacent nodes values based upon the edges connected to them). The most basic adjacency matrix makes use of 0 or 1 to indicate the absence or presence of an edge. However the values in the adjacency matrix can be determined by any score (Wasserman and Faust 1994). Table 3.1 shows an adjacency matrix for 3 nodes in the simple kite diagram:

nodes/nodes	A	B	C
A	0	1	1
B	1	0	0
C	1	0	0

Table 3.1 Adjacency Matrix for nodes A,B and C from the kite network

Although Eigenvector centrality scores can be calculated in a number of ways, in general terms the concept can be expressed mathematically as:

$$x_i = \frac{1}{\lambda} \sum_{j=1}^N a_{i,j} x_j$$

Formula 3.9 Eigenvector Centrality (Craven 2013)

Let x_i denote the score for the i^{th} node and let $A = (a_{i,j})$ be the adjacency matrix of the network. Hence $a_{i,j} = 1$ if the i^{th} node is linked to the j^{th} node, and $a_{i,j} = 0$ otherwise.

Constructing an adjacency matrix of the graph, shown in table 3.1, and solving the eigenvector equation we find the largest eigenvalue to be 3.9423. This gives the corresponding absolute eigenvector [0.3035, 0.4502, 0.2838, 0.4624, 0.2238, 0.3531, 0.4321, 0.2139, 0.0580, 0.0147], where the ordering of the eigenvector corresponds to the

alphabetical ordering of the nodes. Absolute eigenvector in this sense means that all entries of the eigenvector found were negative, and so the values in the eigenvector have been made positive to indicate the relevant sizes of eigenvector centrality scores. The diagram in Figure 3.5 shows the nodes size set to eigenvector scores for the kite network.

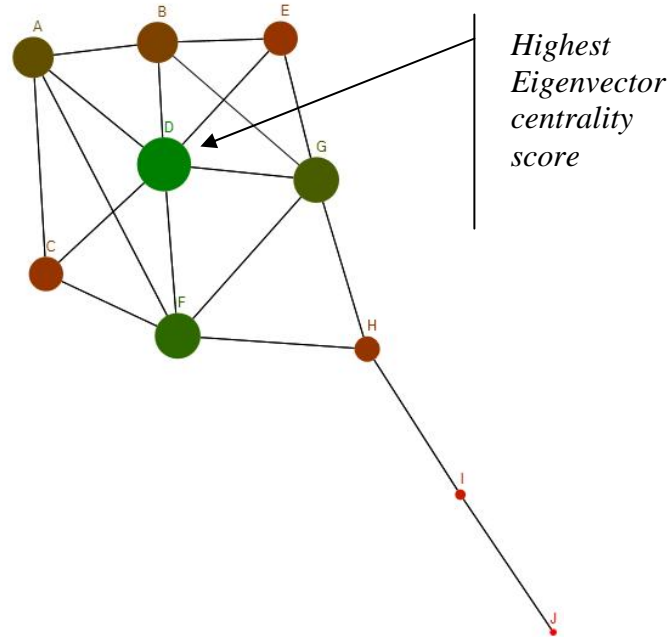


Figure 3.5 Basic graph with node size given by Eigenvector centrality score

In summary, SNA provides a framework of analysis including concepts such as degree, betweenness, closeness and eigenvector centrality scores that have the potential to provide insight into the interactions in social media. The next step is to assess the usefulness of SNA as a medium for measuring social capital.

3.5 Studies using Social Network Analysis in Social Media

There have been a number of studies which examine the social networks underpinning social media which utilise concepts from SNA. These studies are not directly related to the examination of social capital, which will be considered in the next section, but they do provide a rationale for the use of SNA in examining online social graphs. For example

Willging (2005) used degree centrality, betweenness and clique scores in order to analyse a discussion forum and was able to find central members, bridges (betweenness) and social isolates. In addition Willging (2005) argued that SNA was able to uncover relationships not revealed by other analytical methods. Furthermore it was suggested that raw data such as number of posts was likely to overlook structural characteristics of the social graph (Willging 2005).

Mislove et al (2007) examined a number of social media sites including Orkut, YouTube and Flickr using SNA. In Mislove et al (2007) it was argued that the use of SNA would enable an in-depth analysis of such software and afford information likely to be useful in the design of social media systems. The study made use of the SNA concepts of degree (in-degree and out-degree), link symmetry and clustering coefficient. The results of the study indicated that the social graphs were made up of high levels of link symmetry and clusters of low degree nodes connected to other clusters by high degree nodes (bridges). In addition it was found that the graph contained a large, densely connected core and was linked together by about 10% of the nodes with the highest degree. The study also found that path lengths in the core were short (Mislove et al. 2007).

In another study concerned with examining social media by means of SNA Catanese et al (2011) analysed Facebook. The study made use of degree distribution, centrality measures, clustering coefficients and eigenvector centrality scores. It was suggested in the paper that social media should be studied as online interactions would increasingly mirror real world communities and were rapidly becoming the tools of choice for communication. The study found that the higher the degree the lower the clustering coefficient scores. The finding is perhaps to be expected, given the formula for cluster coefficient is the ratio of number of actual links over number of possible links between neighbouring nodes. The study also

examined eigenvector centrality scores and found that they decreased by ranked degree i.e. the lower the rank the lower the eigenvector centrality scores. Again this finding is to be expected as eigenvector centrality scores are based on the scores of linked nodes (Catanese et al. 2011).

In summary, these studies show that SNA can provide a useful analytical tool for understanding the relationships in social media. These studies suggest that the use of SNA metrics such as degree centrality, betweenness centrality, closeness centrality, clustering coefficient and Eigenvector centrality is justified in the study of social media. In addition to these SNA studies of social media there have been studies utilising SNA in the study of social capital, which are explained in the next section.

3.6 Social Network Analysis Measures of Social Capital

Merging the fields of social capital theory and the analysis of social media has face validity. After all, social capital is a measure of social connections and social media software exists to facilitate social connectivity. Furthermore there has been extensive use of Social Network Analysis in understanding social relationships as explained in the previous section. However social capital theory does not rely on connections in a social graph alone. The study of social networks is usually concerned with connections which have been described as social structure with social content (Moody and Paxton 2009). On the other hand the theory of social capital makes use of terms such as bonding and bridging social capital which have the potential to enrich understanding about the nature of social media usage. Therefore it has been suggested that social capital and the analysis of social networks together can provide richer theory and better methods (Baker and Faulkner 2009).

An example of how the combination of social capital and social network analysis can work in practice can be seen in the example of membership of an online community supported by

social media. From the field of social capital theory it is possible to describe the development of trust, norms and shared values. Furthermore we can use the concepts of bridging and bonding social capital to understand how different relationships make up a person's social capital. On the other hand, SNA can be used to examine the interactions that take place between users and therefore suggest how the network structure might support the development of trust and shared values.

In merging the fields of social capital and SNA, it is suggested that SNA provide rigorous concepts and mathematical models. Concepts such as degree centrality, betweenness, clustering coefficient and eigenvector centrality scores provide metrics for social graphs defined by dynamic social interactions, the very mechanisms that are of interest to social capital theorists. For example a triad of nodes is said to be in balance when friends of friends are connected and to be imbalanced when they are not. It has been argued that as actors move to balance their ties, it shapes the overall structure of the graph and therefore affects the flow of information (Davis and Leinhardt 1972; Doreian et al. 1996). Overall, SNA offers social capital theory precise metrics of social graphs, rigorous mathematical models and detailed theories of network formation.

The cross fertilisation between social capital theory and SNA also offers a great deal to SNA researchers. For instance social capital theory and related research can assist in identifying which types of network links are relevant to various social interactions. For example, the measurement of a tie may be dependent on the nature of the relationship. Social capital theory can help to provide the context that shape relationships. For instance the tendency of individuals towards homophily (preference for similarity) in groups has been much studied by social capital theorists (Allport 1954; Schofield 1979; Moody 2001). The theory of homophily can provide context for SNA concepts such as clustering coefficients.

Some theorists have begun to map a relationship between SNA and social capital: for example, in a seminal paper by Borgatti, Jones and Everett (1998) the point was made that the focus of social capital theory has been on substantive issues rather than methodological ones. Furthermore, it was argued that the set of core SNA measures are closest to the description of social capital. The relationship between the SNA measures and social capital as described by Borgatti, Jones and Everett (1998) are laid out in table 3.2.

SNA	Description	Relation to Social Capital
Degree	Number of nodes connected to a given node	Positive - measure of an individual's social capital
Density	Proportion of nodes connected	Negative - if all nodes are tied to each other they are redundant.
Distance	Length of paths across the graph	Positive/Negative- shorter paths mean faster exchange of information
Heterogeneity	The variety of nodes with respect to demographics	Positive - diversity in social capital
Homophily	Ties to similar nodes	Negative- less exposure to a range of ideas although this may lead to faster exchange of information in sub-graphs
Compositional Quality	Number of nodes with required characteristics	Positive - the more potentially useful sources of social capital
Closeness	Graph theoretic distance from a node to all others in the network. An inverse measure of centrality, large values indicate less centrality	Negative - the greater the distance between nodes, the less chance of receiving social capital
Betweenness	The number of times that node fall along the shortest path	Positive - nodes with high betweenness link together sub-graphs otherwise un-connected and thus enabling the flow of social capital
Eigenvector	The extent to which a node is connected to high value nodes	Positive - high eigenvector scores indicate connection to other high social capital sources
Clustering coefficient	Extent to which the graph is divided into cliques	Negative- fractional networks have slower dispersal of information

Table 3.2 Adapted from Borgatti, Jones and Everett (1998) Mapping of SNA to SC

The mapping laid out in table 3.1 illustrates how SNA metrics can be used to measure social capital (Borgatti, Jones and Everett 1998).

The idea of measuring social capital by means of SNA was also examined by Lin (1999) who suggested that as social capital attempts to capture value in social relationships, network

locations should facilitate, but not necessarily determine access to resources such as information. Furthermore it was suggested that requiring high network density or closure to facilitate social capital is not a necessary condition. Research in social capital theory has suggested that bridges are significant in social graphs (Granovetter 1973; Burt 1992). Lin (1999) expanded the discourse on the connection of SNA to social capital to include what was described as cyber-networks. These cyber-networks were essentially early web sites which, it was argued, had the potential to revolutionise social capital (Lin 1999).

These two sources, Borgatti, Jones and Everett (1998) and Lin (1999), are significant as they draw together the prospect of social capital theory enhanced by a rigorous methodological base supplied by SNA. Furthermore Lin (1999) held out the tantalising prospect of cyber-networks supporting social capital. However these interesting treatises on the theoretical linkages between SNA and social capital do not provide validation that SNA can capture the user experience of social capital. It is, however, suggested that the two fields, SNA and social capital theory, can provide a route to a greater understanding of the dynamics of social media.

3.7 Overview of Methods from the Literature

A summary of the research methods and metrics from the literature on social capital measurements in relation to the internet and social media is given in Table 3.3.

Source	Methods	Metric for Social Capital
(Wellman et al. 2001)	Survey	Subject self assessment of network capital, community capital and community commitment
(Hampton 2003)	Survey and ethnographic data	Aid, information and companionship
(Rafaeli, Ravid and Soroka 2004)	Network Analysis	Degree in terms of density of ties
(Wasko and Faraj 2005)	Network Analysis Survey	Degree centrality
(Best and Krueger 2006)	Survey	Trust, reciprocity and integrity

(Zhao 2006)	Survey	Degree: number of friends
(Williams 2006)	Survey	Created ISCS including bonding and bridging social capital
(Smith 2008)	Network Analysis	IAN and ESN correspond to the weak and strong ties
(Ellison, Steinfield and Lampe 2006) (Ellison, Steinfield and Lampe 2007) (Steinfeld, Ellison and Lampe 2008)	Survey	FIS for bonding and bridging social capital
(Valenzuela 2008; Valenzuela, Park and Kee 2009)	Survey	Life satisfaction, social trust, civic participation, political participation.
(Pfeil, Arjan and Zaphiris 2009)	Network Analysis	Number of friends (degree)
(Gilbert and Karahalios 2009)	Network Analysis and survey	Tie strength (degree)
(Burke, Kraut and Marlow 2011)	Survey and server logs	ISCS for bonding and bridging social capital
(Ellison, Steinfield and Lampe 2011)	Survey	FIS for bonding and bridging social capital
(Yoder, Hill and Stutzman 2011)	Survey	FIS bridging social capital via (Ellison, Steinfield and Lampe 2007)
(Gil de Zúñiga, Jung and Valenzuela 2012)	Survey	Six item survey based upon Lin (2008) e.g. 'feel connected'
(Antoci, Sabatini and Sodini 2012)	Network Analysis	Ties in an online network (degree) through time.
(Hofer and Aubert 2013)	Survey	ISCS for bonding and bridging social capital
(Jung et al. 2013)	Experiment using a survey	ISCS for bonding and bridging social capital and further subdivisions
(Ellison et al. 2014)	Survey	Adapted ISCS
(Sun and Shang 2014)	Survey	Items developed from (Nahapiet and Ghoshal 1998) use social interaction ties, shared vision and trust

Table 3.3 Summary of Methods and Metrics for social capital from the literature

The table 3.2 shows that the most common method of measuring social capital in social media in the literature is the survey. The dominance of the survey method is likely to have been because it is the most direct means of obtaining a subject's experience of social capital. It is also evident that the ISCS (Williams 2006) is the most influential single measure of social capital. The survey method provides the advantage of gaining insight into the user's experience of the software beyond the realm of simply counting interactions or understanding relative positions in the social graph. On the other hand self-reporting of user experiences

can lead to misleading results particularly where users may be embarrassed to report accurate information.

In addition to surveys there have been a number of studies utilising Social Network Analysis (SNA) methods most particularly degree centrality. However in some case the SNA derived measures have been theoretical and lacked supporting data (Smith 2008; Antoci, Sabatini and Sodini 2012). In other cases the SNA measure has been degree centrality mapped to number of friends or ties in the network (Rafaeli, Ravid and Soroka 2004; Wasko and Faraj 2005; Pfeil, Arjan and Zaphiris 2009, p. 643–654; Gilbert and Karahalios 2009). In addition the SNA measures previously used were not systematically validated against a recognised measure of social capital such as the ISCS. This leaves open the possibility that these approaches are not accurately measuring social capital.

Given this gap in existing body of knowledge the aim of this research involved creating SNA based metrics that was validated against ISCS data. This approach was adopted in order to ensure that the metrics were accurately measuring social capital. The next section contains an overview of the findings from the literature review

3.8 Overview of the Literature in Social Capital Measures in Social Media

The purpose of this literature review was to determine **precedents**, **omissions**, and **guidance** which can assist in the creation of social capital metrics. This review has also highlighted **key findings** in the literature that can be addressed by the use of new social capital metrics.

There are ample **precedents** for the use of social capital theory as applied in a social media context dating back over a decade up to the present day (Wellman and Frank 2001; Best and Krueger 2006; Gil de Zúñiga, Jung and Valenzuela 2012; Ellison et al. 2014). Whilst it is clear that surveys are often used to measure social capital, there have been a number of

studies making use of SNA particularly in terms of degree centrality (Rafaeli, Ravid and Soroka 2004; Wasko and Faraj 2005).

There are **omissions**, or gaps, in the existing knowledge in terms of the lack of a valid SNA based formula postulated to measure social capital (Rafaeli, Ravid and Soroka 2004; Smith 2008). It is clear that there has been an understandable dominance in the use of surveys to measure social capital (Zhao 2006; Steinfield, Ellison and Lampe 2008; Valenzuela 2008; Yoder, Hill and Stutzman 2011; Gil de Zúñiga, Jung and Valenzuela 2012). However it is equally clear that surveys cannot produce an automated metric. In summary there is a lack of valid SNA based metrics which has been validated as measuring social capital.

The literature provides **guidance** in terms of containing the highly influential ISCS (Williams 2006) which has been used to measure social capital in a range of studies (Ellison, Steinfield and Lampe 2006; Burke, Kraut and Marlow 2011; Hofer and Aubert 2013; Jung et al. 2013). It is also worth highlighting the fact that ISCS (Williams 2006) makes use of Putnam's influential ideas of bonding and bridging social capital (Putnam 1995, 2000). The literature concerning the use of SNA in the measurement of social capital also contains the idea that SNA measures such as clustering coefficient may make a contribution (Borgatti, Jones and Everett 1998).

The process of developing valid metrics for social capital outlined in this research also included the application of the metrics to two separate case studies, as will be explained below. Using the proposed new metrics enables this research to cast light on **key findings** in the literature. For example Steinfield, Ellison and Lampe (2009) found that amongst users of social media, bonding social capital increased over time. A finding that is examined below. However some of the findings in the literature are disputed. For example two studies in the literature suggested that social capital can be measured by number of friends (Zhao 2006;

Pfeil, Arjan and Zaphiris 2009). Whilst on the other hand Ellison, Steinfield and Lampe (2011) suggested that this was true only up to a point. Furthermore in a more recent study Ellison et al (2014) asserted that social capital was not generated by data such as friends but rather by interactions such as postings (Ellison et al. 2014). The key findings outlined here are discussed below in the light of evidence found using the proposed new metrics.

3.9 Chapter Summary

In this chapter the literature concerning measuring social capital in social media has been discussed. It has been shown that there are two major research methods used i.e. the survey, and SNA mainly using degree centrality. Furthermore the Internet Social Capital Scale survey has been influential in measuring social capital. The next chapter contains an outline of the preliminary analysis carried out for the practical work in this thesis.

Chapter 4 Preliminary Analysis of Social Media Usage

4.1 Overview

This chapter contains an explanation of the results from a preliminary analysis of a case study of social media usage. The preliminary analysis was carried out to determine if Facebook would be a suitable case study for the development of social metrics and related software. The preliminary analysis was made up of interviews, surveys and social network analysis the results of which are discussed below.

4.2 Preliminary Analysis: Student Facebook Groups

When setting out on a voyage of discovery a degree of preparation usually pays dividends. Therefore before commencing with the development of metrics, let alone the related software, it was thought prudent to carry out a preliminary analysis. The preliminary analysis was used to determine if a Facebook group used by students in an academic setting was a suitable test bed for a wider investigation into developing social capital metrics. In order to carry out this preliminary analysis a number of commonly used research methods were deployed. Interviews and a survey were used to determine how students used three example student Facebook groups. Finally Social Network Analysis (SNA) measurements were used to examine the structure of the underlying social graphs of the Facebook groups.

4.3 Interviews and survey

The research questions examined were firstly ‘what were the Facebook groups being used to discuss’, and ‘what were student attitudes to using the group’. These questions were addressed by means of interviews with individual users; the results of these interviews were used to devise a question set for a survey.

The interviews were carried out with the six student volunteers; each paid £10.00 in Amazon gift vouchers. The interviews were semi-structured with set questions and free flowing discussion. The topics discussed during the interviews were:

- The nature of the group
- Contributions to the group
- Social aspects of the group
- Informational aspects of the group

These topics of discussion yielded the following findings

- The group was used mainly for subject information and occasionally arranging social gatherings.
- The group was followed usually at least once a day
- Contributions in terms of postings were occasional because participants felt that they didn't have anything to add or felt that they might sound stupid
- The social discourse and opportunities to speak to different year's cohorts was welcome
- The information in the posting was considered valuable

These findings from the interview (given in appendix B) were used to create a survey also given in appendix B. The sample group for the survey was made up of the thirty student volunteers, each of whom were entered into a prize draw for £20.00 of Amazon gift vouchers. The key findings are discussed below.

Question\Score	Strongly Disagree		Neutral	Strongly Agree	
	Disagree	Disagree		Agree	Agree
The course FB group helped with my general knowledge about EH	0.00%	0.00%	19.40%	64.50%	16.10%
The course FB group helped me find specific answers to questions	3.20%	12.90%	29.00%	45.20%	9.70%
The course FB group helped me make new friends	6.50%	19.40%	41.90%	29.00%	3.20%

Table 4.1 Survey results of student use of the course Facebook group

The results in table 4.1 indicate that over 80% of students ‘agree’ or ‘strongly agree’ that the group was useful for general subject knowledge. Over half the students ‘agree’ or ‘strongly agree’ that it was useful for specific answers to questions. Around 30% ‘agree’ or ‘strongly agree’ that it was useful for making friends.

The interview responses indicated that there was reluctance on the part of users to post to the group; this is a finding that was echoed in the survey results. Over a quarter of the respondents either agreed or strongly agreed that the group was ‘intimidating’. However over half of the students didn’t find the group intimidating as shown in table 4.2 below. It was also found that a majority of students found the group ‘welcoming’, ‘informative’ and ‘entertaining’.

	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
Welcoming	9.7% (3)	61.3% (19)	25.8% (8)	3.2% (1)	0.0% (0)
Intimidating	3.2% (1)	25.8% (8)	16.1% (5)	45.2% (14)	9.7% (3)
Informative	35.5% (11)	58.1% (18)	3.2% (1)	3.2% (1)	0.0% (0)
Entertaining	25.8% (8)	58.1% (18)	12.9% (4)	3.2% (1)	0.0% (0)

Table 4.2 Survey results of student attitudes to the course Facebook group

Taken together, the results from the interviews and survey response indicated that the Facebook group has a dual purpose of both providing information and for building social ties. This was a significant finding as a major aim of the preliminary analysis was to determine if

the software could provide a test bed for a study into both bonding and bridging social capital. Therefore these initial results were supportive of using this social media context as a test bed for developing social capital metrics.

4.4 Social Network Analysis

Social Network Analysis (SNA) was carried out on the postings of three Facebook groups because it provided a means of investigating the structural relationships of the social graph in the groups. The SNA measures used were: degree centrality, betweenness centrality, closeness centrality, clustering coefficient and eigenvector centrality.

The posters and commenters were extracted from the group via the Facebook Application Program Interface (API). There were three data sets taken from three separate academic Facebook groups covering a twelve week time period. The three data sets related were A and B relating to two courses related groups and C relating to one module related group. The posters were arranged in one column of data with commenters on the posts arranged in a second column of data. The two column data set was then imported into NodeXL which was used to calculate the various SNA metrics and to produce graphs. NodeXL is a plugin for MS Excel which expands the available tool set and provides a means of graphing the data (Nodexl.com 2014). There are other SNA software tools such as Gephi, UNICET and Pajek but these were rejected in favour of NodeXL as it is was considered to be the most flexible tool for computing additional metrics and tailoring intuitive to read graphs. The results of the SNA analysis of the three data sets are given in the following sub-sections.

4.4.1 Degree centrality

Degree centrality is a traffic measurement where users are nodes and the interactions between users are the edges. In a directed graph for the data sets, the in-degree is messages received and the out-degree is messages posted. In directed graphs degree centrality is therefore the

sum of in-degree and out-degree. In case of this analysis, out-degree is a post to anyone who comments and in-degree is comments to the post.

Data set	Minimum	Maximum	Mean	Median
A	0	38	10.305	7.00
B	0	72	9.78	6.0
C	0	34	4.604	3.00

Table 4.3 Degree Centrality Results

The results in table 4.3 show for data set A the range of postings was 0-38. The results indicate that for data set B there was a larger range in number of posts i.e. 0-72 and a mean of 9.78. The data set C is in respect of a Facebook group set up for discussions concerning an individual module and it was found that the mean and median number of posts is relatively low in comparison to the other two data sets which were set up to service entire undergraduate courses. The degree centrality data for data set A was used to construct a graph as shown in figure 4.1.

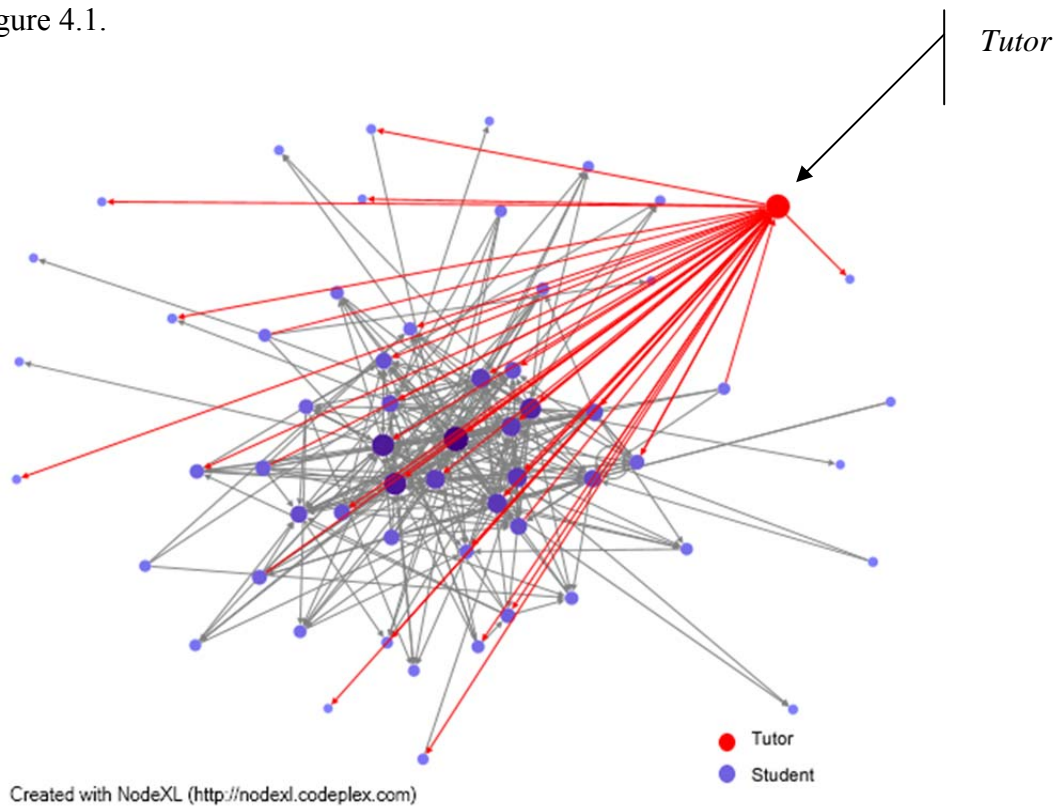


Figure 4.1 Data Set A 'degree centrality' score indicated by node size

The figure 4.1 above shows a tutor indicated in red in the graph with a largest node indicating that they had the highest degree centrality. As previously explained the graph is directed by post as out-degree and comments as in-degree. In reading the graph you should take into account that the larger the node is the greater the number of posts/comments (or higher total degree). With the exception of the red node which is highlighted together with its edges to illustrate the degree centrality of the largest node. The graph as a whole is low density 0.08 (high density would be everyone connected to everyone else). There are sixty nodes and 315 edges; duplicates are not shown. The graph illustrates that there is a wide range in the size of the nodes. It is worth noting that 'lurkers', i.e. users who merely view the postings, are not shown in this graph.

In a sense the graph in Figure 4.1 is a way of visualising the frequency of posts. However it is worth noting that examining degree centrality scores in this manner also reveals the structural relationships of the social graph.

4.4.2 Betweenness Centrality

The graphs showing betweenness scores for each of the three data sets are shown in figures 4.2, 4.3 and 4.4. The betweenness centrality, i.e. the number of times a node is on a shortest path; in this context is essentially a measure of the number of different people with posts that are commented upon. If a person commented on posts from everyone they would have the maximum betweenness score. In the figures below the larger the betweenness score the larger the node shown as circles. In all three graphs the tutor's circle is in the top right hand corner. In comparison to the tutor other users have low betweenness scores (shown as smaller circles), this shows that most users commenting on only a few users posts.

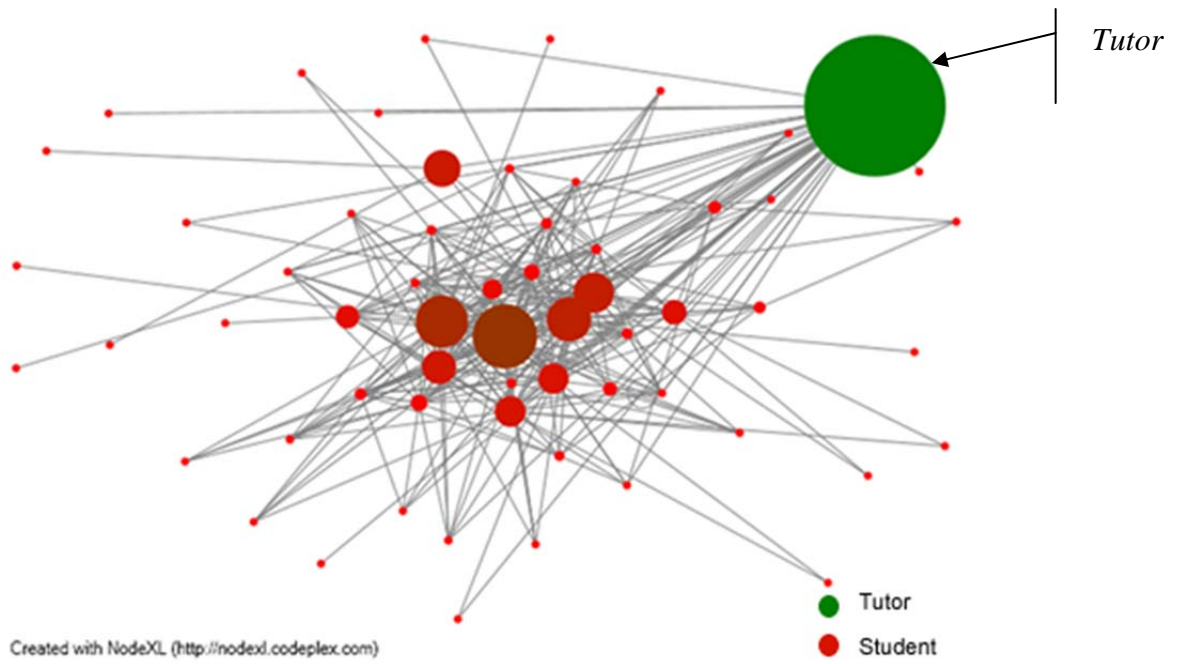


Figure 4.2 Data Set A Betweenness score indicated by node size (tutor top right)

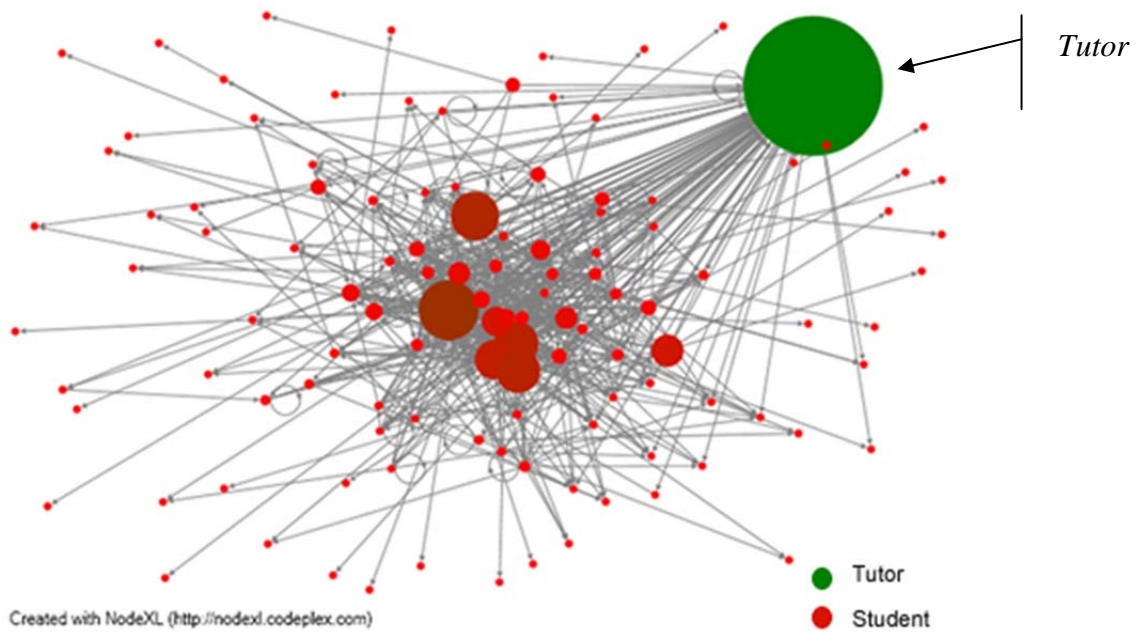


Figure 4.3 Data Set B Betweenness score indicated by node size (tutor top right)

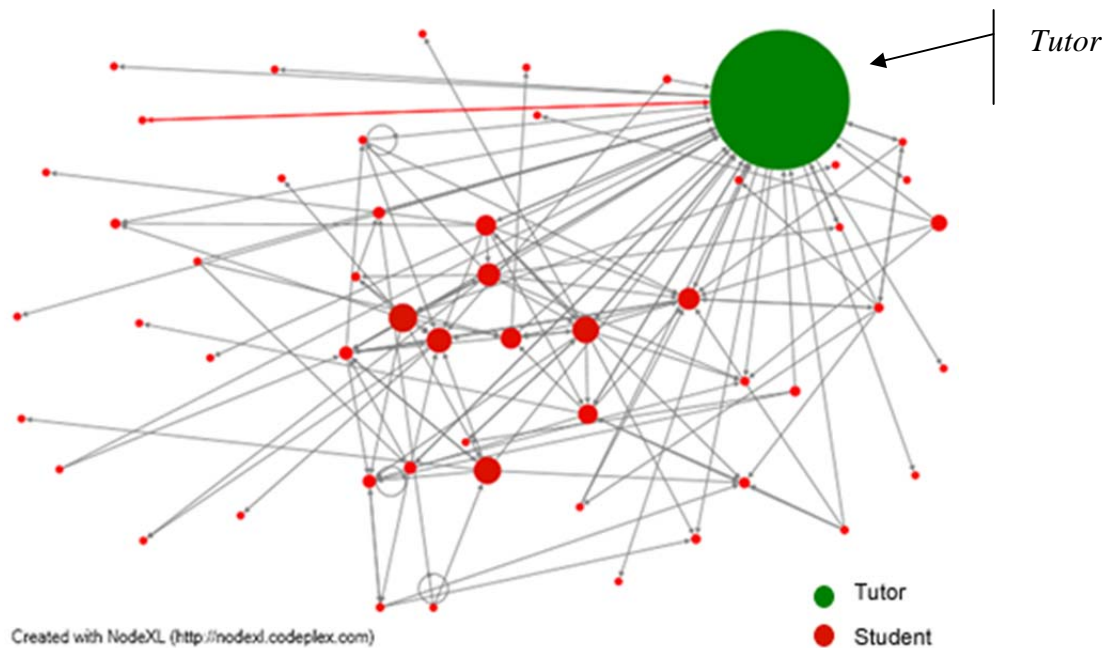


Figure 4.4 Data Set C Betweenness score indicated by node size (tutor top right)

The graphs in figures 4.2, 4.3 and 4.4 show a similar pattern. The fact that the three graphs have a similar pattern is significant for this preliminary analysis because it indicates that Facebook groups of this type have characteristics in common. If there are common characteristics this in turn suggests that results for a group may be generalisable to other Facebook groups.

In reading the graphs it should be noted that the largest node in each of the graphs represents a lecturer as opposed to the other nodes which are all students. The graphs are illustrating that there are a lot of nodes with low betweenness centrality scores, and in comparison to the tutor's node there is not a large range in the scores. These graphs indicate that the tutors did not discriminate with whom they communicated with, and they communicate relatively frequently, as in fact one might hope. On the other hand student users are more discerning about the posts that they choose to comment upon, and the comments on their posts are from a relatively small number of fellow students.

4.4.3 Clustering Coefficient

Clustering coefficient is a basically a measure of cliqueness or the propensity of people to interact with a small group of people. It is calculated by arranging connected nodes into triads. A hit represents a triad where nodes are connected in a triangle; a miss represents a triad with only two connecting sides. Clustering coefficient is calculated as hits divided by hits and misses. The graphs showing clustering coefficient scores for each of the three data sets are shown in figures 4.5, 4.6 and 4.7. The higher the clustering coefficient scores the larger the node.

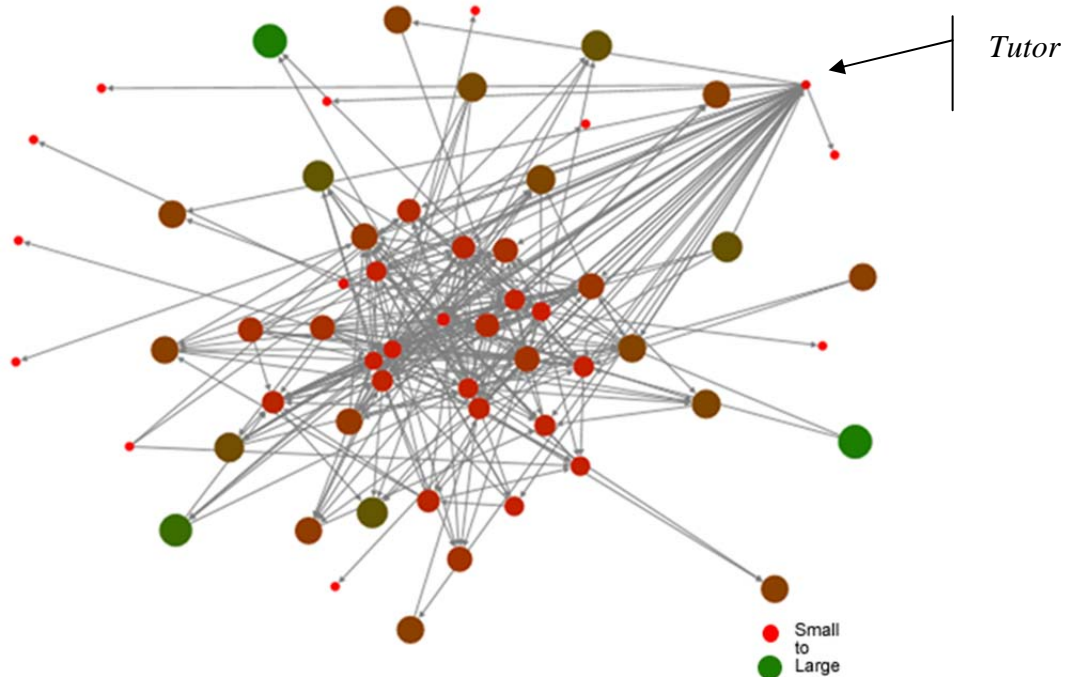


Figure 4.5 Data Set A clustering coefficient score indicated by node size

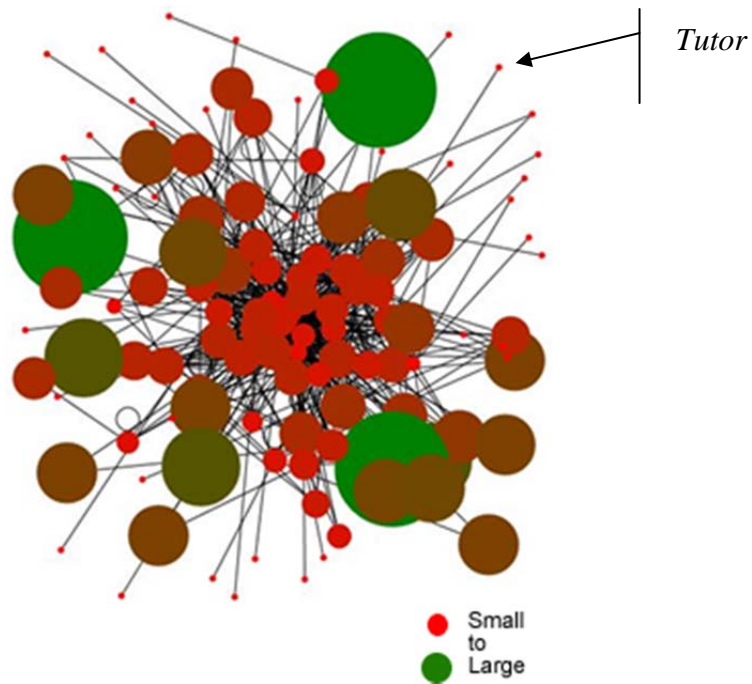


Figure 4.6 *Dat Set B clustering co-efficient score indicated by node size*

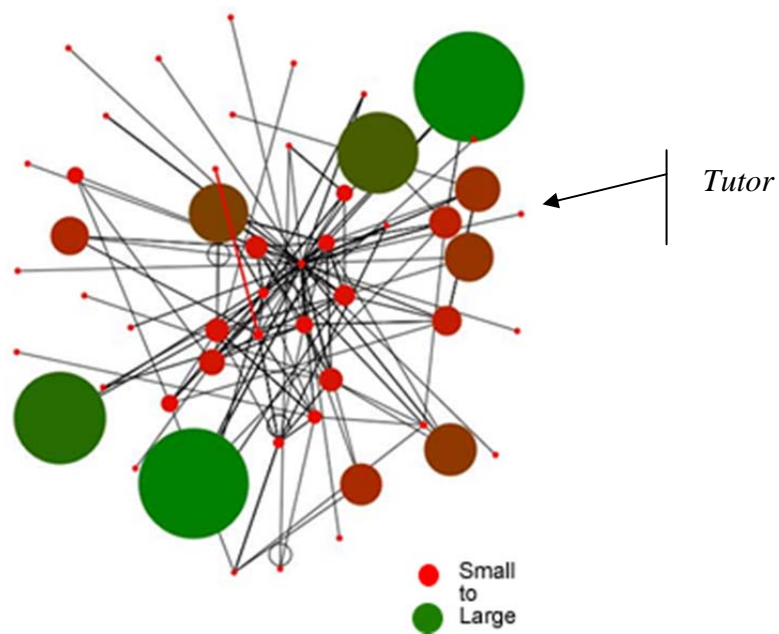


Figure 4.7 *Data Set C clustering coefficient score indicated by node size*

Note firstly that Figures 4.5, 4.6 and 4.7 graphs each have a high number of relatively large nodes, in other words users with high cluster coefficient scores. This is particularly striking

in comparison to the betweenness and closeness graphs shown above. Put simply the clustering coefficient graphs, figures 4.5, 4.6 and 4.7 have a greater number of larger circles than those for betweenness shown in figures 4.5, 4.6 and 4.7. However these results are consistent with the finding from the betweenness scores, in so far as they indicate that students are selective about the posts they choose to comment upon.

In summary there are several interesting aspects to the clustering coefficient graphs, firstly the fact that users are showing a preference in whom they choose to message. This characteristic might relate to bonding social capital as it describes connections to friends. Secondly the very fact that there are wide variations in scores at least indicates a difference in the experience of using the software. Finally the similarity in the three graphs is further evidence that these Facebook group would make a consistent case study.

4.4.4 Eigenvector Centrality

The eigenvector centrality value is concerned with popularity. It is a metric based upon first scoring nodes, and then sharing out the scores amongst the nodes connected on the graph. Therefore a node scores more highly by being linked to a ‘popular’ node than to a ‘loner’ node. The results of the Eigenvector centrality for the three data sets are given in table 4.4.

Data set	Minimum	Maximum	Mean	Median
A	0	0.052	0.017	0.014
B	0	0.044	0.008	0.005
C	0	0.084	0.019	0.014

Table 4.4 Eigenvector centrality values across the 3 data sets

The results in table 4.4 show that data set B with the smallest number of contributors has the lowest mean and median Eigenvector centrality score of 0.008. This result reflects the fact that data-set B has a lower number of ‘popular’ nodes therefore the start values for the eigenvector centrality scores were lower.

In addition to the Eigenvector centrality scores, Pagerank scores were also calculated for each data set. Although Pagerank is a variant of the eigenvector centrality formula (Page et al. 1999), there is a variation in the dampening factor also known as the scaling factor. In essence Eigenvector centrality scores can be calculated through a number of iterations i.e. you calculate the new score and carry out the sharing process over and over again. In the absence of a dampening factor, an eigenvector centrality score will grow exponentially at each iteration of the calculation.

Pagerank by Eigenvector x,y charts were plotted for all three datasets and are shown below in Figures 4.8, 4.9 and 4.10. Pagerank values were plotted along the y axis, Eigenvector values along the x axis, with the degree centrality score of each node indicated by size.

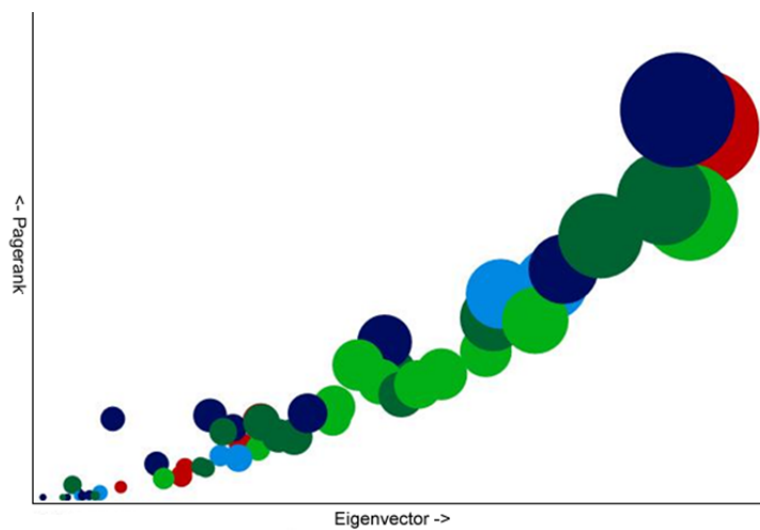


Figure 4.8 Data Set A Pagerank (y-axis) Eigenvector (x-axis) degree centrality by size

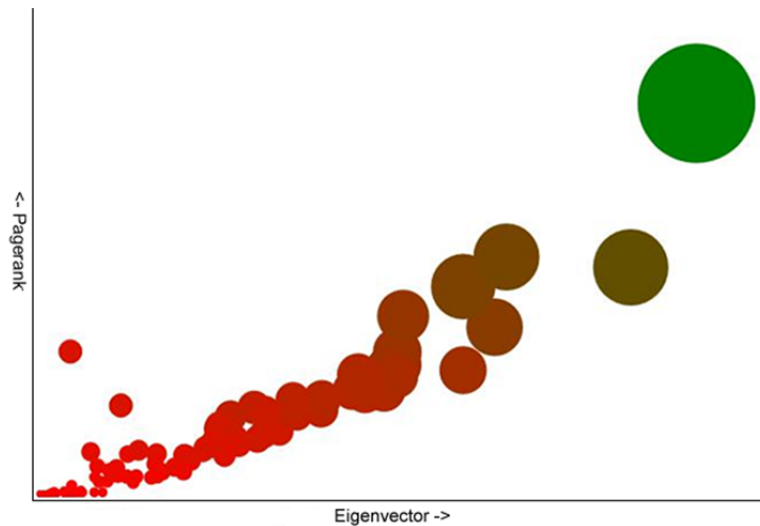


Figure 4.9 Data Set B PageRank (y-axis) Eigenvector (x-axis) degree centrality by size

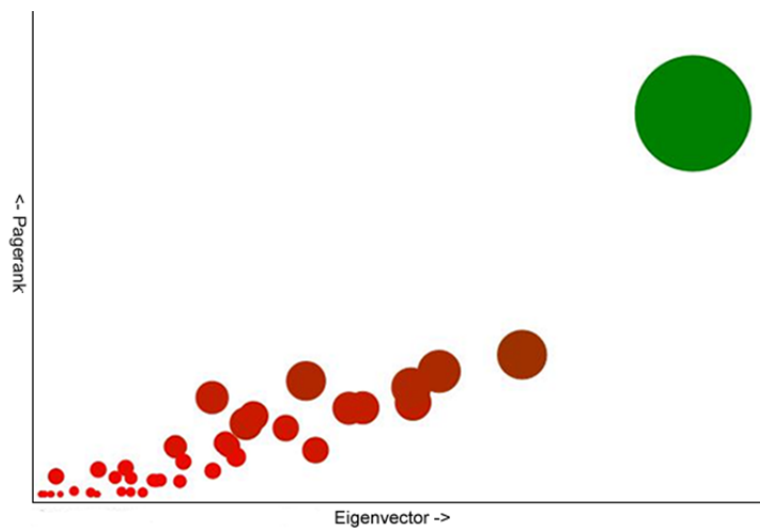


Figure 4.10 Data Set C PageRank (y-axis) Eigenvector (x-axis) degree centrality by size

In the process of interpreting figures 4.8, 4.9 and 4.10 it is useful to note that if PageRank and Eigenvector values were the same for each node the data points on the chart would be in a straight line; furthermore that if nodes with a high degree score (they are the larger circles) also had high eigenvector centrality/PageRank scores you would expect to see smaller nodes in the bottom left of the chart progressing to larger nodes in the top right of the chart.

In respect to the eigenvector centrality scores and Pagerank scores there is a general tendency in the plotted chart, shown in the figures, towards a straight line but that this is not absolute. In other words, as might be expected, the scores are similar but not the same. In terms of the of high degree nodes being in the top right of the charts, again it can be concluded that there is a tendency for high scoring nodes in degree to also be high scoring nodes in eigenvector centrality and Pagerank.

In summary there are interesting aspects to the graphs in figures 4.8, 4.9 and 4.10. Firstly the results show that there is a wide range in eigenvector centrality scores and this bears a relationship to the amount of posting, or degree centrality.

4.4.5 Content Analysis

The primary focus of this research is on measuring the value of social relationships in social media. This focus on the structural connections in a social graph means that the content of messages is not important. However, for completeness, one of the data sets was analysed for content. The Facebook group examined was centred on a single module. It had the smallest data set and therefore was used as an example of how content analysis could be applied.

Each of the messages was categorised by subject e.g. 'assessment' or 'lecture' and so on.

Degree centrality and Eigenvector centrality scores were calculated and shown in figure 4.11 and 4.12 respectively.

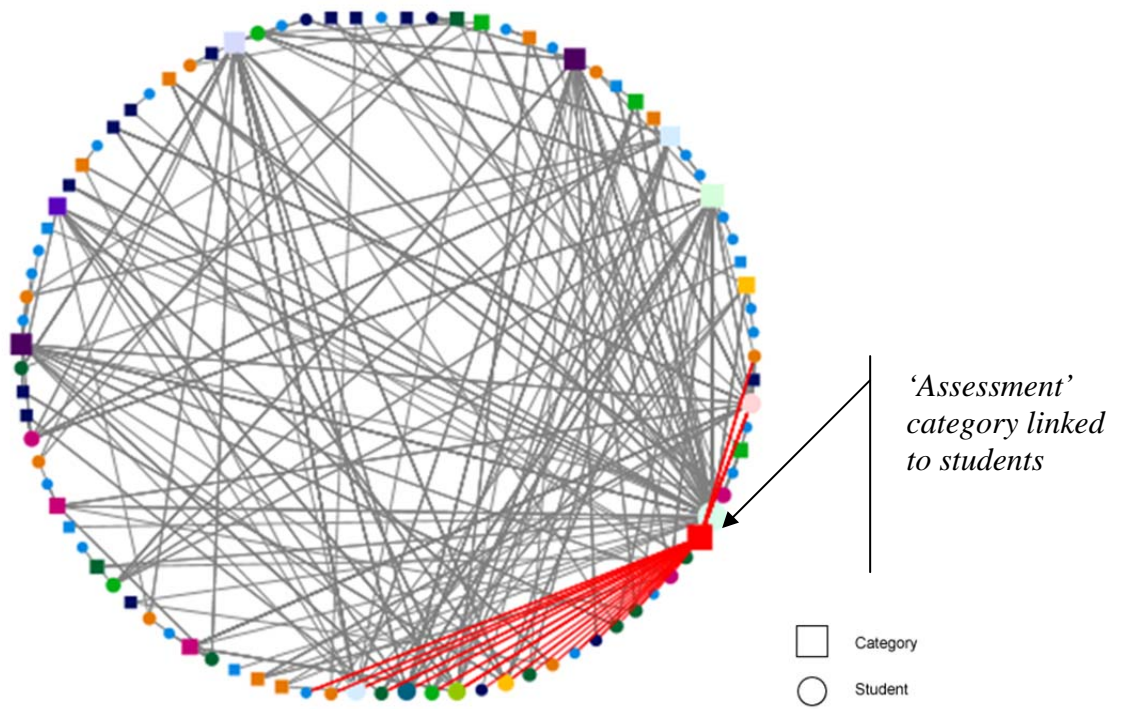


Figure 4.11 Data Set C Content Analysis by Degree Centrality

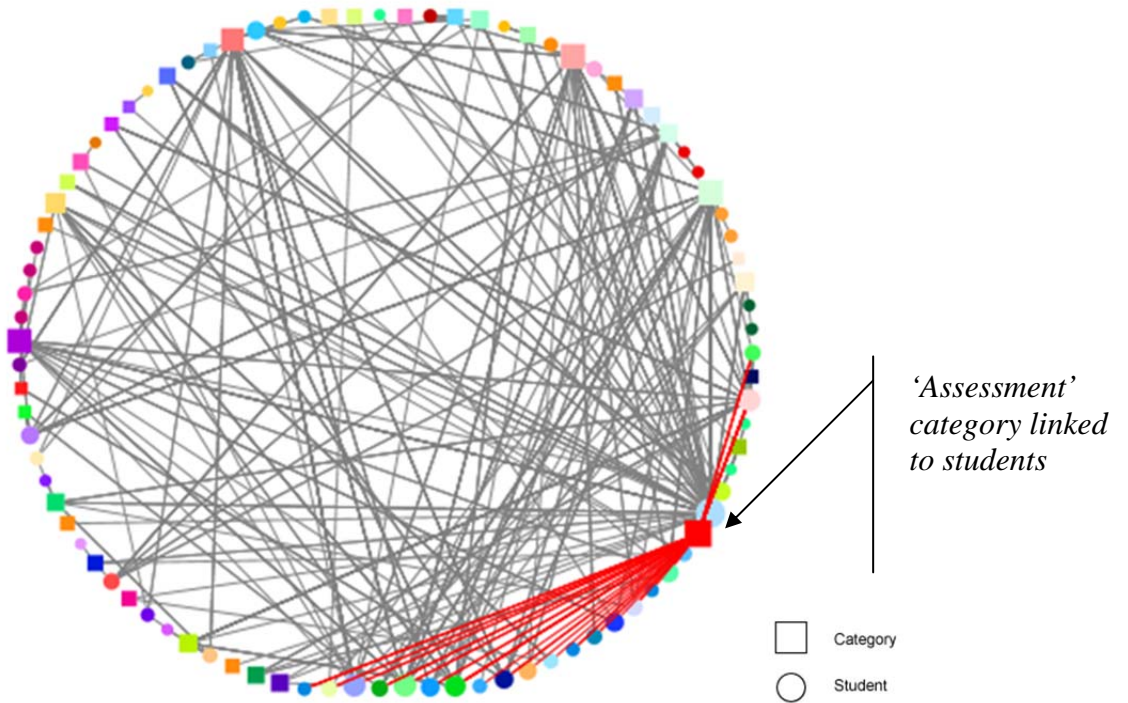


Figure 4.12 Data Set C Content Analysis by Eigenvector Centrality

The graphs in figures 4.11 and 4.12 show the categorised messages as a square, for example the square highlighted in red in both graphs is for 'assessment'. The 'assessment' square represents all messages with that categorisation. The circles in the graph represent contributors, either posters or commenters. This strand of investigation demonstrates the outcome of SNA when used for content analysis. The graphs show the complexity of the relationships between users and message based upon the content. The findings of such an investigation can be interesting but from a software engineering perspective the logical outcome will always be algorithms that categorises the messages. It is suggested that whilst this type of research may have value in the broader area of social media research it is not related to the central purpose of this study which is to measure social capital in social media.

4.5 Discussion

The intention of the preliminary analysis of students' Facebook groups was to determine if this social media context was appropriate as a test bed for social capital metrics. Furthermore it was intended to uncover some of the underlying properties regarding the use of the software.

It was found by means of user interviews and a survey that the Facebook group made a valued addition to the students' experience and that the social media software was providing access to both friendships and knowledge. This duality of function provided by this case study of social media usage is particularly important as it is an intention to examine bonding and bridging social capital in arriving at social capital metrics.

The results indicated by the SNA showed that there were a relatively few users with high degree centrality scores. The results also suggested that there were relatively few users with high eigenvector centrality scores and in addition those users typically had low betweenness scores but high clustering coefficient scores. These results, concerning clustering coefficient

held out the intriguing possibility that users were tending to post to friends, which may in turn indicate that the software could yield a valid measure of bonding social capital. Moving forward to the main focus of this research, these findings suggested that degree centrality and clustering coefficient looked particularly promising as metrics for social capital.

The SNA results yielded similar patterns of results across all three data sets e.g. the clustering coefficient scores. This suggests a degree of consistency that indicates that a student Facebook group would make a useful test bed for social capital metrics.

In summary there were a number of indications which suggested that the software could yield a valid measurement for social capital and that SNA could provide insights into the structural relationships in the students' social media graph.

4.6 Ethical Research Procedures

The conduct of ethical research does not solely lie in the adherence to a narrow set of rules or procedures. Ethics in research should emanate from a general standpoint or philosophy about how research should be conducted. The ethical framework for this research was broadly based upon the four moral principles of autonomy, non-maleficence, beneficence and justice laid out by Beauchamp and Childress (2001). The principle of autonomy refers to the right an individual has to decide to participate in a study or not. The concept of autonomy is linked to the ideas of veracity (honesty), fidelity (trust), confidentiality and privacy. In this study the principle of autonomy was used to inform the practice of recruitment of volunteers and in gaining informed consent. Non-maleficence relates to the idea of doing no harm to participants in the study. In this study the principle of non-maleficence was used in judging the scope of the work. For example asking subjects about their broader personal experience of social capital could be potentially intrusive; this is why the study focuses on the much narrower idea of social capital in relation to the use of a particular instance of the use of

social media. The concept of beneficence refers to the research being useful to the research community and to society as a whole, whilst the principle of justice refers to treating all participants in a study fairly and equitably. In this study the principles of beneficence and justice were characterised by the overarching aim of the study which is to improve the functionality of social media software in such a way as to contribute to the research community.

In the addition to the adherence to an ethical framework there were a number of ethical issues relevant to research in the area of social media in particular. It has been suggested that because social media data is public that there is no necessity to seek ethics approval (Solberg 2010). Furthermore there has been a published Facebook study where user privacy consent was not sought (Lewis et al. 2008). However on the other side of the argument it has been suggested that there is a growing awareness of the importance of ethical use of social media data (Henderson, Hutton and Mcneilly 2012), and in addition it has been suggested that informed consent should be obtained on a case by case basis (Moreno, Fost and Christakis 2008). It is also worth noting that there are issues regarding the use of social media data related to the platform provider and in addition there are legal issues in terms of the requirements of the Data Protection Act.

It is possible to extract data from someone's social media account without their knowledge or consent. However using such data is not only questionable ethically, in our view, it is poor research practice. The problems of this approach starts with determining the sample group. Which collection of presumably un-willing participants do you select? Furthermore, in the case of this study, it begs the question what do you use as a means of determining a reliable measure of their social capital. Relying on number of friends or followers in a social media site is only a reliable measure of social capital 'up to a point' as we have seen in the literature

review. Given the issues of ethics and reliability it was decided that the use of a student Facebook Group would offer an avenue for investigation. It is notoriously difficult to ensure subject anonymity in social media research, but by keeping both the subjects names private and also the name of the group private it is unlikely that subject identity will be revealed. Using a Facebook group also meant that it was possible to create a group which clearly stated that the interactions in the group were subject to analysis and thus inform users that research was under way. As a practical matter it also made it possible to recruit users to participate in a study.

The steps taken in order to create an ethical framework include:

- Abertay ethics approval for all methods
- Informed consent of participants completing: preliminary surveys, ISCS survey and FIS survey, use of Facebook data
- Research usage of Facebook groups a link to information about the study
- Compliance with Facebook terms and conditions as at 2013-14
- Subject data was anonymised to the fullest extent possible.

4.7 Chapter Summary

This chapter contained an explanation of the research methods used to carry out a preliminary analysis of student social media usage. The results with particular importance are summarised as follows:

- Interviews – the Facebook group was valued as a source of knowledge and friends.
- Survey – a majority of students agreed that the Facebook group provide them with general subject knowledge, specific answers to questions. Around a third of participants found the software helped them to make friends.
- Social Network Analysis (SNA) – degree centrality provided a means of identifying the structural relationships in the social graph. Betweenness and Clustering coefficient

scores suggested that students preferred to communicate with users who were identifiable as cliques of friends.

- The results of SNA across three different Facebook groups showed discernible patterns.
- The results of a content analysis using SNA were presented but it was suggested that content analysis is not directly related to social capital.

In the light of these findings it is argued that a student Facebook group is a suitable test bed for the development of social metrics and related software. The fact that the software has a dual purpose of providing information and access to friends suggests that it may map successfully to the concepts of bridging and bonding social capital discussed above.

Furthermore support for this prospect was found by the fact that SNA measurements degree and clustering coefficient were particularly interesting metrics. The fact that similar patterns emerged between the three Facebook groups suggests that results from one group may be generalisable to other groups of this type. Based on the findings in this preliminary analysis the work on validating and evaluating social capital metrics was carried out, and is explained in the next chapter.

Chapter 5 The SCiSM Metrics

5.1 Overview

The aim of this research is to create valid metrics of social capital that can be utilised in measuring the performance of social media software. This chapter contains an explanation of the practical work carried out in pursuit of the aim.

The **proposed** metrics styled as Social Capital in Social Media (SCiSM) is outlined in three mathematical formulae relating to bonding social capital, bridging social capital and total social capital.

A key goal of this research is to **implement** the metrics in software firstly in order to illustrate a proof of concept and secondly to make the metrics available to other researchers. Therefore this chapter contains a description of the software which was written in php using Model View Controller (MVC) architecture. This approach was used in order to make the software usable and extendable.

The chapter also contains the results of a study outlining the **validation** of the SCiSM metrics using correlations and linear regression against the influential Internet Social Capital Survey (ISCS) scores proposed by Williams (2006). The SCiSM metrics were validated against two independent data sets which were new to this research i.e separate from the data sets used in the preliminary analysis. The Meneely (2012) validation methodology was utilised. For the purposes of experimental control the metrics were also examined against the Facebook Intensity Scale (FIS) (Ellison, Steinfield and Lampe 2007) in order to eliminate the possibility that the metrics were simply measuring Facebook use rather than social capital per se.

In developing new metrics it is crucial that the measurement is an improvement on existing measures. Therefore this chapter contains a study outlining an **evaluation** of the SCiSM metrics against other Social Network Analysis (SNA) measures. The SNA measures used in the comparative evaluation included degree centrality, clustering coefficient and eigenvector centrality.

The primary purpose of developing new metrics is to be able to apply them in a range of useful contexts. Therefore in order to **demonstrate** the utility of the SCiSM metrics, two data sets were analysed and the results related to key findings from the literature.

5.2 Methodology for Validating the Metrics

It is essential to ensure that the validation criteria are not arbitrary but rather the criteria used are supported by an appropriate philosophy. In the case of this research the methodology for validating software metrics proposed by Meneely (2012) was used. In practice using the Meneely (2012) methodology involved assessing the key aims of the metrics and from these aims deriving criteria which in turn were used to construct the methods.

When using the Meneely (2012) methodology the first step in validating a metric is to consider how the metric is intended to be used. In the case of the SCiSM metrics it is suggested that the metrics could be used to evaluate the performance of the software. The next step of the Meneely (2012) methodology is to look up an ‘advantage’ appropriate to the intended use. In this case the advantage of ‘efficiency’ was considered to be of central importance to the metrics because one of the intended goals of the SCiSM metrics was to assess the social media software’s efficiency in terms of fitness for purpose. The next step involved using a table of criteria provided by the Meneely (2012) methodology in order to determine which criteria were associated with the ‘efficiency’ advantage. The Meneely (2012) methodology states that the efficiency advantage mapped to ‘improvement validity’

(over other existing measures) and ‘usability’ (feasible to measure). Therefore the first two Meneely recommended criteria are:

- Improvement validity
- Usability

In addition to these two validity criteria the Meneely (2012) methodology states that the following criteria are also applicable:

- Association – a direct statistical correlation with an external factor (Fenton 1994; Rao 2007)
- Empirical validity – experimentation corroborates the relationship between the metric and an external factor (Briand, El Emam and Morasca 1995; Kitchenham, Pfleeger and Fenton 1995)
- Predictability – predict values of an external factor (Fenton 1994; Roche 1994)
- Underlying theory validity – based upon an underlying theory that has validity in the domain of the application (El Emam 2000; Kitchenham, Pfleeger and Fenton 1995)

The additional criteria were added because they broaden the base for ensuring the SCiSM metrics were valid.

The validity criteria were then mapped to specific methods as shown in the following table.

Criteria	Methods (location in thesis)
Underlying theory	Mathematical formula for the proposed metrics (section 5.3)
Usability (feasibility)	Implementation in software (section 5.4)
Association Empirical validity Predictability	Correlation between metrics and an existing valid measure (section 5.5)
Improvement validity	Comparison of scores with other metrics (section 5.6)

Table 5.1 Validity criteria mapped to research methods

It is suggested that the above represents a robust methodology for validating a social capital metric.

5.3 Proposed Social Capital In Social Media (SCiSM) Metrics

In the process of proposing original metrics for social capital, we must first address the question ‘Why create a new measure?’. The rationale for devising new metrics is based upon an analysis of the literature. From the literature, summarised above, it is clear that the measurement of Social Capital has relied heavily on surveys. Surveys such as the ISCS (Williams 2006) are extremely useful in arriving at an understanding of how people perceive the elements of social capital in social media. However survey-based metrics suffer from the limitations of self reporting and do not take account of actual interactions with the software. Furthermore, adopting the survey approach is not useful in developing software that measures software performance as this research requires. Put simply, social capital as measured by surveys can’t be automated in a way that is scalable.

The literature also contains studies which put forward mathematical models (Smith 2008; Rafaeli, Ravid and Soroka 2004) of how to measure social capital but these studies do not provide a validation of the models based upon actual social media usage data. In other words these are proposed formulae without the process of validation against data to support the proposition. It is therefore argued, based upon the findings in the literature review, that there is a case for new social capital metrics.

The next question to be addressed is how best to construct such metrics. The literature suggests that SNA provides a rigorous set of techniques which can be deployed to examine relationships in a social graph. The literature also contains studies demonstrating the applicability of SNA for the study of social capital. However, it is worth noting that new social capital metrics must measure a property of a social graph in a way that is not measured by the traditional SNA techniques.

A further question which needs to be addressed in the process of developing original social capital metrics is ‘What does a new measure have to accomplish?’. Social capital metrics must take account of the network structure (social graph) and the different affordances of the software, such as messaging, commenting and viewing. It is suggested that the metrics must also take account of strong and weak ties in the network structure, representing bonding and bridging social capital respectively. In addition the metrics must also be a valid measure of social capital.

For convenience in referring to the proposed metrics it will be styled as Social Capital in Social Media (SCiSM) pronounced ‘schism’.

5.3.1 Bonding SCiSM

The first task at hand is to model a formula that embodies bonding social capital. Bonding social capital, in essence, refers to an individual’s close (or strong) ties such as with friends or family members. The process of proving the formula involved calculating subject scores which were then validated against ISCS scores. However the validation of the metrics against a particular data set cannot be extrapolated to all other data sets. Therefore it is suggested that is important that the metrics are based upon an existing metric which has a logical underpinning supporting a prime facia case for believing that the metric is measuring social capital. In the case of bonding SCiSM the formula is derived from the SNA measure clustering coefficient sometimes referred to as a clique score. The clustering coefficient score is adapted in bonding SCiSM by the introduction of additional factor which is the subject’s out-degree. The SNA measure of clustering coefficient was selected as it is a measure of connectivity of nodes and was therefore thought likely to be translatable to bonding social capital. The subject’s out-degree was added to the metric because implicit in the idea of bonding is that an action is needed to make a bond. For example if you post a

message you may not even read all the comments (in-degree) therefore there is no certainty that bonding social capital exists, however in posting a message (out-degree) there is an act of communication which can be reasonably modelled as bonding social capital. All of the SCiSM scores, including bonding SCiSM, are bracketed in time segments.

Thus the logical underpinning for the bonding SCiSM scores is that it models a subject's level of cliqueing, taking into account their active participation, in time segments. The notion of time segments is important in the development of a metric which is intended to be dynamic and therefore potentially used during the run time of software.

Therefore it is suggested that bonding SCiSM can be measured as follows:

$$\text{bonding social capital}(i, t) = do(i, t) * \left(\frac{\lambda_G(i, t)}{\tau_G(i, t)} \right)$$

Formula 5.1 Bonding SCiSM

Let $do(i, t)$ be the out-degree of posts seen in the specified time segment t for node i . In addition let $\lambda_G(i, t)$ be the number of sub graphs of G with 3 edges and 3 nodes, one of which is i , (essentially hits) whilst $\tau_G(i, t)$ is the number of subgraphs with 2 edges and 3 nodes one of which is i plus the number of subgraphs of G with 3 edges and 3 nodes (essentially misses + hits) in the connected component to 1 degree in the specified time segment. Conceptually this states that bonding social capital is made up of a subject's out degree multiplied by their clustering co-efficient in a time segment. The formula is shown as a visualisation below in figure 5.1.

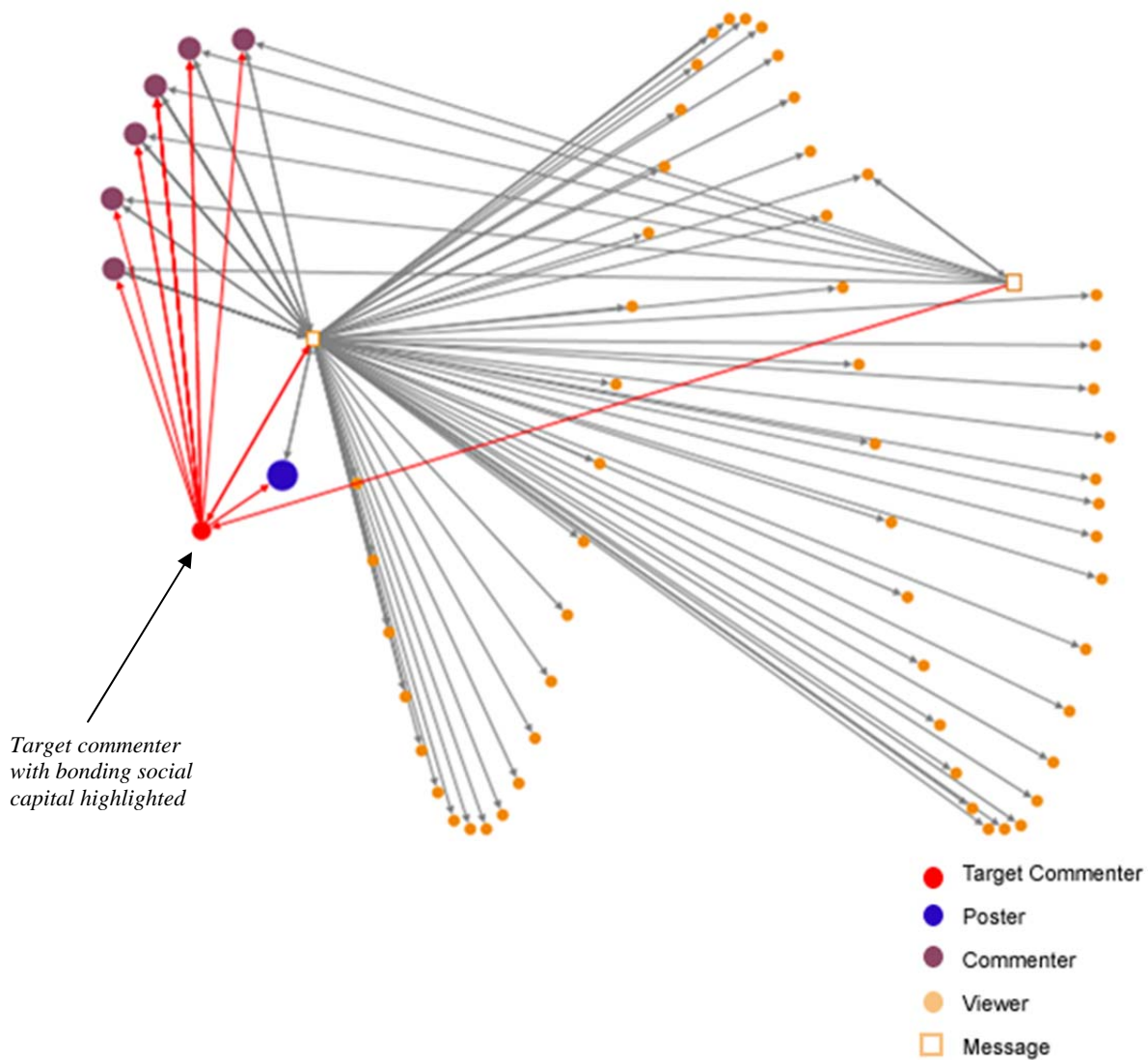


Figure 5.1 Visualisation of a graph for an individual's Bonding SCiSM

Figure 5.1 shows a one week time period with two messages posted shown as squares in the diagram (to the left and right). The node representing the poster of the message is on the left of the graph and shown as the largest blue circle. The purple nodes in the top left corner of the graph are users who commented on the post. The red node, on the left towards the bottom, is another commenter to the message. It is red node's bonding social capital that is shown as red lines. The grey edges represent other postings and viewings. In essence the graph, specifically the red node and red edges, provides a visualisation of the bonding SCiSM formula for one subject. The calculation for the target commenter is as follows:

=9 (out-degree) multiplied by (9 hits divided by 0 misses plus 9 hits)

=9* (1)

= 9

Note that the bonding SCiSM visualisation shown in figure 5.1 encapsulates a situation where the poster and the commenters were the same for both messages, a different circumstance for calculating bonding social capital is given below where total social capital is discussed.

An explanation of the bridging SCiSM formula is given in the next section.

5.3.2 Bridging SCiSM

The next task in modelling social capital was to devise a mathematical formula for bridging SCiSM. The notion of bridging social capital, according to social capital theorists, is the network of acquaintances or weak ties in a social network. It can be thought of as the network structure which provides information flows to and from a subject in a social graph. In the process of devising a credible metric for bridging social capital the SNA measures for in-degree and out-degree were used as these are essentially traffic scores within the social network. Therefore an individual's bridging SCiSM takes account of the actual in-degree (or traffic of information in) from the message and the messages out-degree to all other users in a time segment. In effect the bridging SCiSM score is measuring a subject's share of the traffic in the social network. You may recall that previous researchers have used degree centrality as a measurement of social capital. Therefore we suggest that bridging SCiSM has an underlying logic suggesting that it is social capital that is being measured. What is new in this work is the contextualising the degree centrality score in the message to which it belongs and in taking account of time segments.

Therefore it is suggested that bridging SCiSM can be measured as follows:

$$\text{bridging social capital } (i, t) = \left(\frac{da(i, t)}{dp(i, t)} \right)$$

Formula 5.2 Bridging SCiSM

Let $da(i, t)$ be the actual in-degree of an individual node summed, whilst $dp(i, t)$ is the potential in-degree for the time period for the time segment t . Conceptually this states that bridging social capital is made up of a subject's actual in-degree divided by potential out-degree. The formula is shown as a visualisation below in figure 5.2.

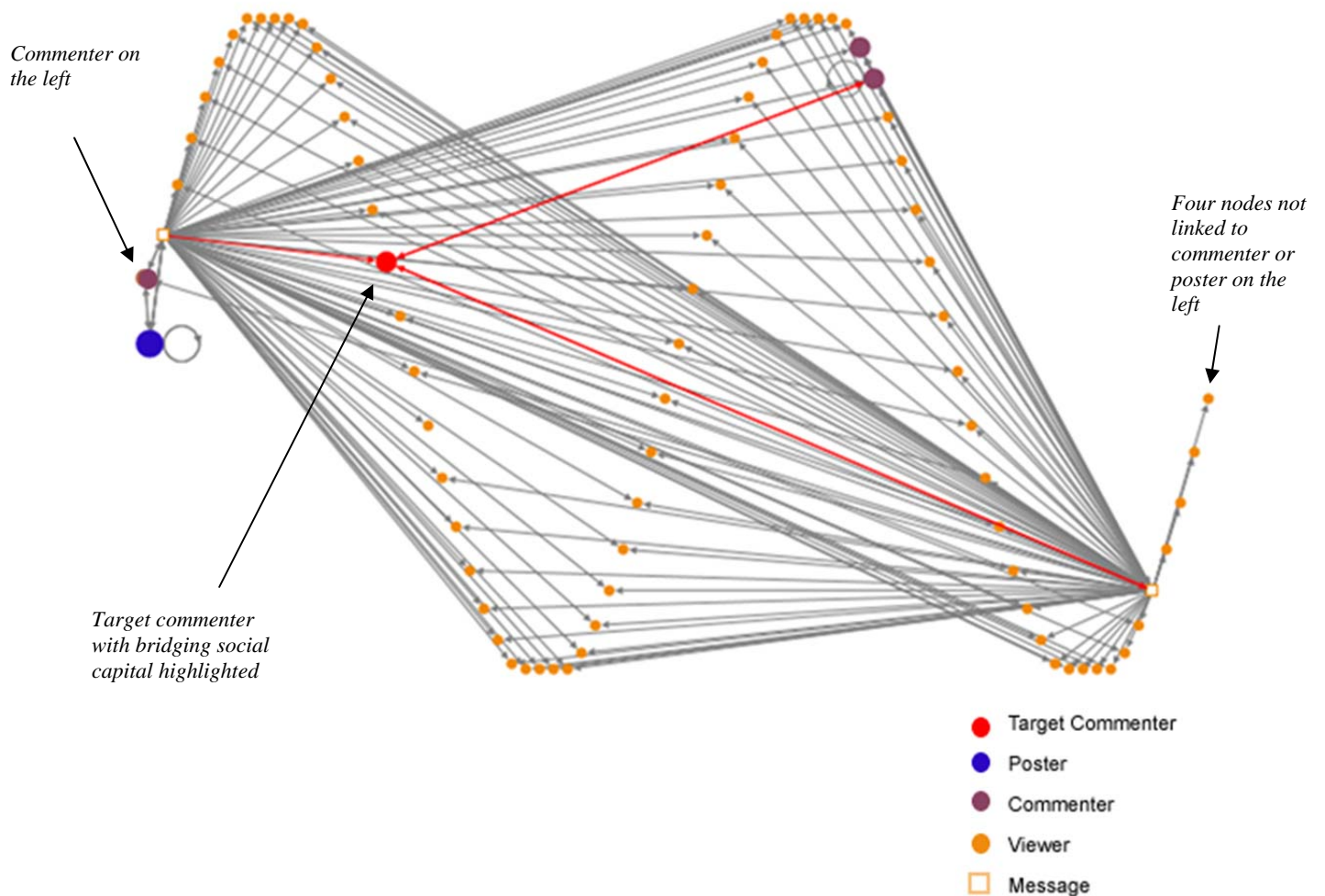


Figure 5.2 Visualisation of an individual's Bridging SCiSM

Figure 5.2 shows a one week time period with two messages posted shown as squares in the diagram (to the left and right). The poster of the message on the left of the graph is shown as the largest blue node. The poster of the message on the right is the lower one of the two purple nodes. The red node is the subject of the highlighted portion of the graph. The red edges indicate the bridging social capital of the red node. The red node viewed both messages and commented on the message to the right. Therefore the red node received in-degree from the two messages as the user saw the message. Incidentally all of the small orange nodes are viewers but not posters. Recall that, in addition, the red node commented on the post on the right and received a comment from the poster. Therefore the red node together with the red edges shows the user's bridging SCSIM. In essence the graph provides a visualisation of the bridging SCiSM formula for the red node. The calculation for the target commenter is as follows:

$$= 3 \text{ (actual)}$$

$$= 75 \text{ (possible)}$$

$$= 3/75 = 0.04$$

For the purposes of contrast the calculation for the commenter, annotated in figure 5.2 on the extreme left, does not include the four nodes on the extreme right and is as follows:

$$= 2 \text{ (actual)}$$

$$= 71 \text{ (possible)}$$

$$= 2/71 = 0.028$$

5.3.3 Total SCiSM

The final task in mathematically modelling social capital is to consider both bonding and bridging social capital together. The process of devising a total SCiSM metric simply sums both bonding and bridging social capital. The underpinning logic is derived from social capital theory.

Therefore it is suggested that total SCiSM can be measured as follows:

$$\text{total social capital}(i, t) = \text{bonding social capital} + \text{bridging social capital}$$

Formula 5.3 Total SCiSM

The formula 5.3 is an amalgamation of the formulae for bonding and bridging social capital.

The formula is shown as a visualisation below in figure 5.3.

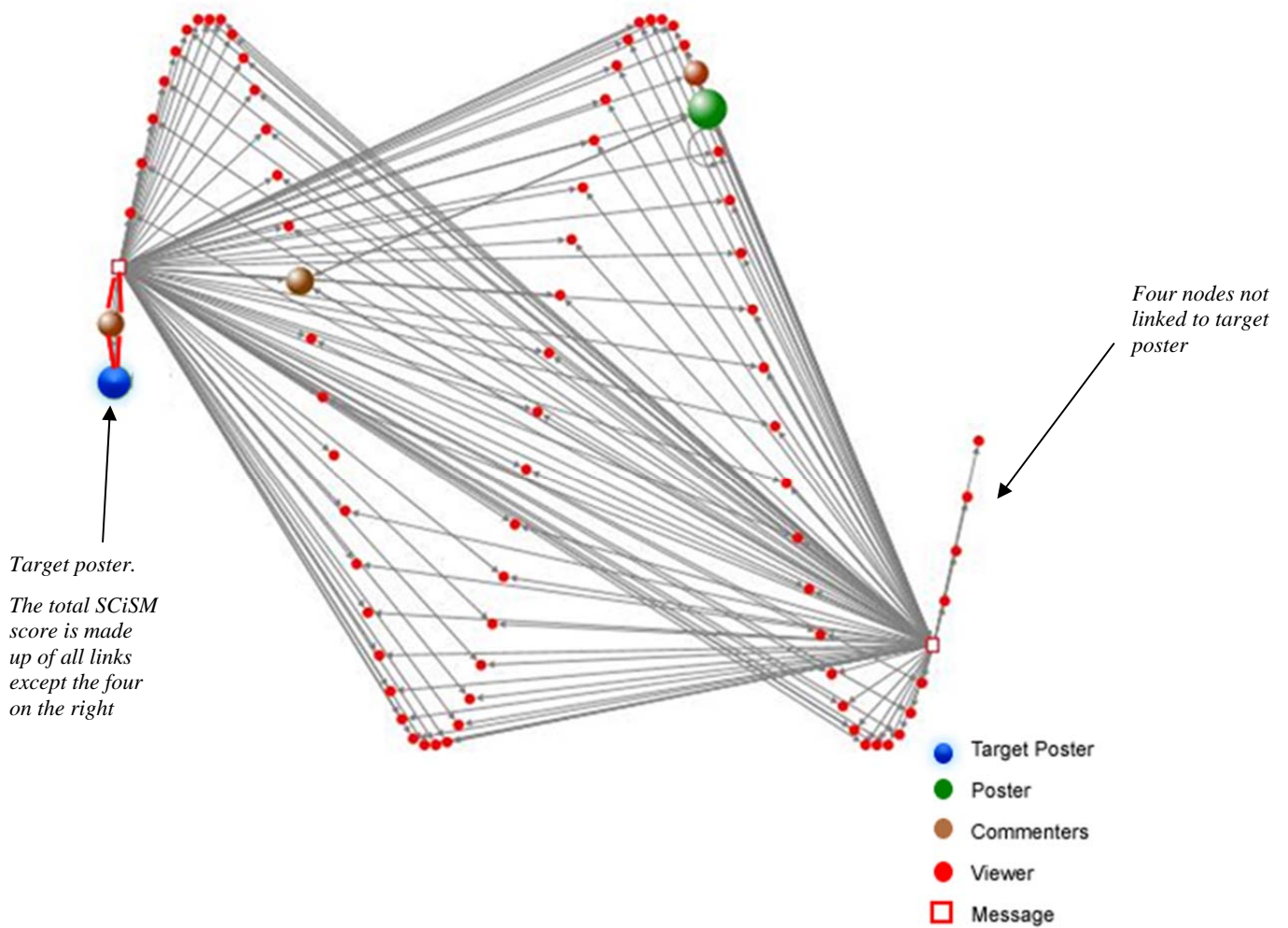


Figure 5.3 Visualisation of graph for total SCiSM

The figure 5.3 above shows the weak tie network edges (bridging social capital) in grey and the strong tie network edges (bonding social capital) in red. The posted message that is ‘seen’ by other nodes is represented by a square, whilst the small red circles are nodes which have seen the message. The blue and green spheres represent nodes which posted messages whilst the three brown spheres are nodes which commented upon the posted messages. In essence the graph provides a visualisation of the total SCiSM scores for the target poster shown in blue.

Bonding SCiSM for the target blue node is calculated as follows:

=2 (out-degree) multiplied by (1 hit divided by 1 misses plus 1 hit)

=2* (1/2)

= 1

Bridging SCiSM for the target blue node is calculated as follows:

= 2 (actual)

= 71 (possible)

=2/71 = 0.028

Therefore the total SCiSM score is the sum of both bridging and bonding SCiSM as is as follows:

Total 2.028

At the centre of these original formulae is the proposition that bridging social capital is the structure for information flow and bonding social capital is the sharing of the information flow in a sub-group. In figure 5.3 the blue lines show bonding social capital, the grey lines show the social information network. It is suggested that the software is providing two distinct functions and that examining these both independently and together, using SCiSM, provides an avenue for analysing the performance of social media software. The next step was to implement the metrics in software.

5.3.4 SCiSM Formula Comparison to Formula in Literature

Above, in the literature review, two formulae for social capital were examined. These were Rafaeli, Ravid and Soroka (2004) and Smith (2008). To reiterate, the Rafaeli, Ravid and

Soroka (2004) formulas referred to group and individual density which it was suggested related to social capital. The SCiSM formulas differ from the Rafaeli, Ravid and Soroka (2004) approach because SCiSM is intended to provide an individual measurement of social capital. The SCiSM metrics are, in the Bourdieu (1985) tradition, not presumed to be of equal applicability for group measurement of social capital. Furthermore in the case of Rafaeli, Ravid and Soroka (2004) in the formula for individual social capital, the degree centrality score is presumed to be made up of reciprocal links and is not articulated as being separated into bonding and bridging social capital unlike the SCiSM metrics.

To reiterate, the Smith (2008) formulas were divided into bonding and bridging social capital as are the SCiSM metrics. However Smith (2008) is concerned with using the attribute of affinity or in other words to consider common features between nodes in the social graph. Therefore Smith (2008) is measuring homogeneity. This is not the approach adopted with the SCiSM metrics which are focused upon the interactions which make up the social graph. It is also worth re-iterating that neither Rafaeli, Ravid and Soroka (2004) nor Smith (2008) presented validation for their metrics.

5.4 Implementing SCiSM in a re-usable extensible software framework

The intention behind implementing the SCiSM metrics in code is to provide a proof of concept that a social capital theory based metrics can be captured in software.

The Model, View, Controller (MVC) software design pattern has the advantage of dividing the view from the implementation details, an approach which is particularly useful given the complexities of user interface presentation. In addition the division of the view from other aspects of the implementation protects the core functionality from debugging updates to the interface. Another important advantage of the MVC pattern is that it enables the development of re-usable and extendible code. On the other hand the MVC pattern has the

disadvantage of not being as intuitively clear. It was decided that the revised MVC pattern represented the best route for creating the SCiSM software intended to contribute to the work of other researchers.

In the original MVC the controller is the starting point to the trio of MVC components. The request is first passed to the controller, which then instantiates the models and views that are required in order to respond to a request from the user. The original MVC architecture specifies that the controller contains the business logic and, and controls how the application responds to user interactions (Reenskaug 2003). The original MVC architecture is outlined in figure 5.4.

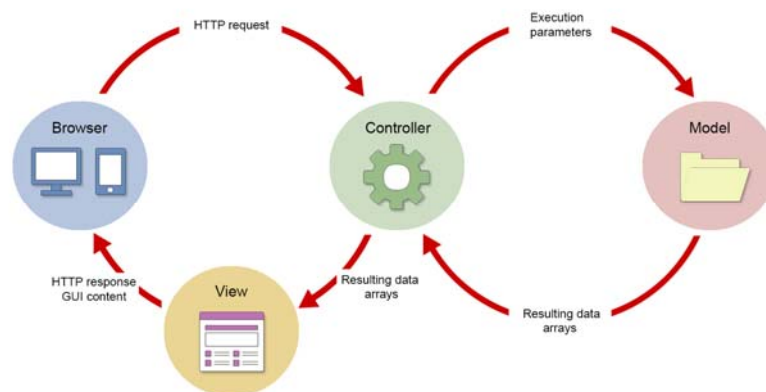


Figure 5.4 Original MVC architecture

There have been subsequent revisions to the original pattern as the original MVC architecture suffers from a limitation when it is implemented for a web based application. The limitation arises from the fact that the view is dominant in web application development in part due to the HTTP protocol but also stemming from the user interface design which is heavily focused on response to user actions. Therefore several sources use a revised form of the MVC architecture (Butler 2012; Garfield 2006) in which the role of the view is more prominent as illustrated in figure 5.5.

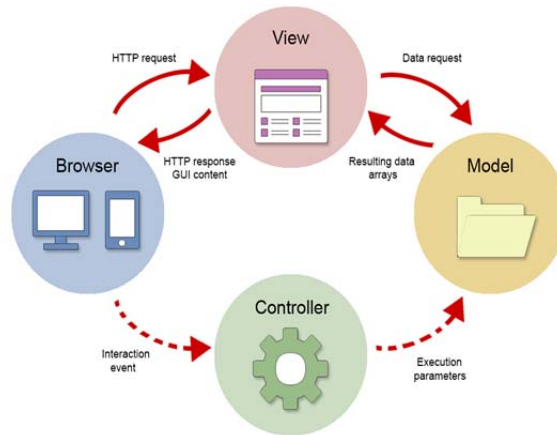


Figure 5.5 Revised MVC architecture used for SCiSM

In summary the revised MVC architecture illustrated in figure 5.5 shows the arrangement whereby a HTTP request is responded to initially by either the controller (or the view) which in turns executes the methods in the model in order to create results which are then parsed to the view component. The revised architecture was used for the SCiSM metrics.

The scripting language selected was PHP primarily because of its popularity. In addition a search was conducted for existing SNA libraries in server-side languages in July 2014. There were two major software libraries found for the python programming i.e ‘libsna’ (at <http://www.libsna.org/>) and ‘snap’ (at <http://snap.stanford.edu/links.html>). However no SNA libraries written in PHP were found. It was concluded that given the popularity of PHP implementing the metrics in the language would provided a useful contribution to other researchers. Furthermore the PHP server-side programming language supports the creation of classes and is also amenable to using the MVC design pattern.

5.4.1 SCiSM Software Architecture

The SCiSM software architecture is illustrated in figure 5.6.

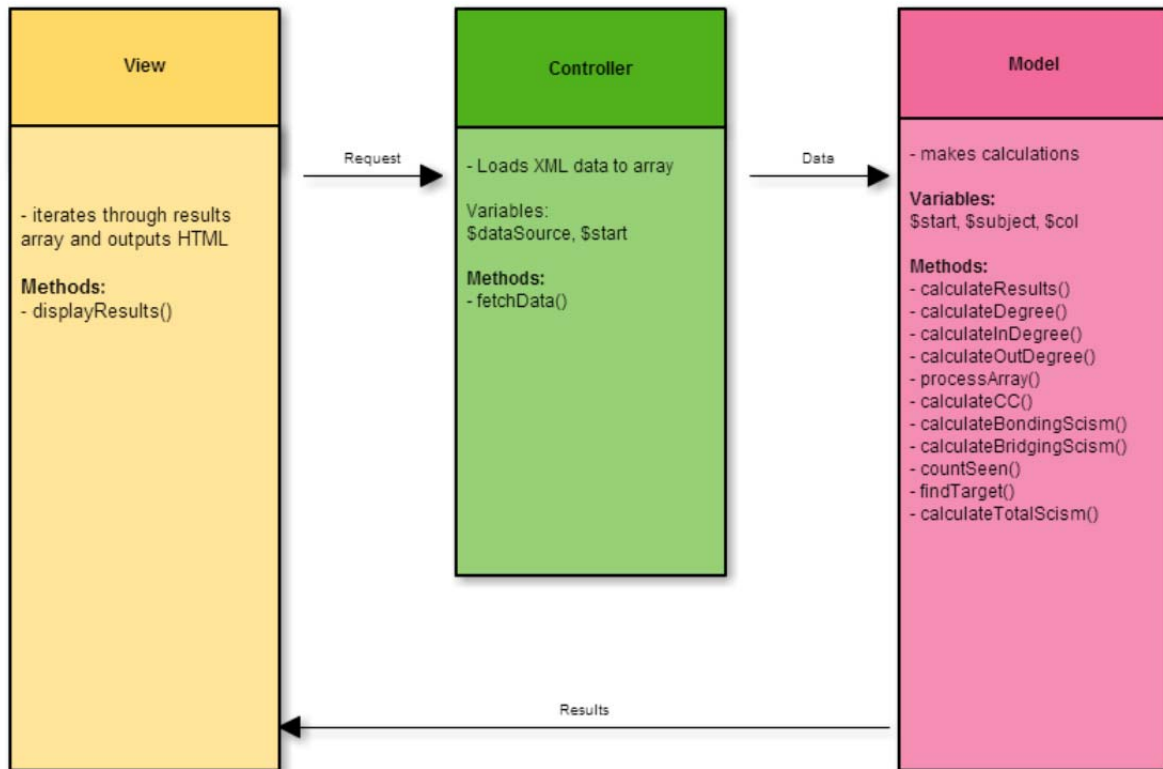


Figure 5.6 SCiSM software architecture

The diagram above shows the SCiSM software marked out in an MVC framework with the intention of providing reusable and extensible software. The view element of the framework is simply concerned with displaying the results in HTML. The controller is responsible for loading the data. The model element of the framework is responsible for the data calculations. In order to add functionality to the software another developer would alter the data handling in the controller and add the calculation functionality to the model.

5.4.2 SCiSM Software & Web Site

The code embodying the SCiSM metrics is given in appendix C. In order to make available the SCiSM metrics to other researchers the functioning software is available via a web site at - <http://sociallearning.space.abertay.ac.uk/scism/> shown in figure 5.7.

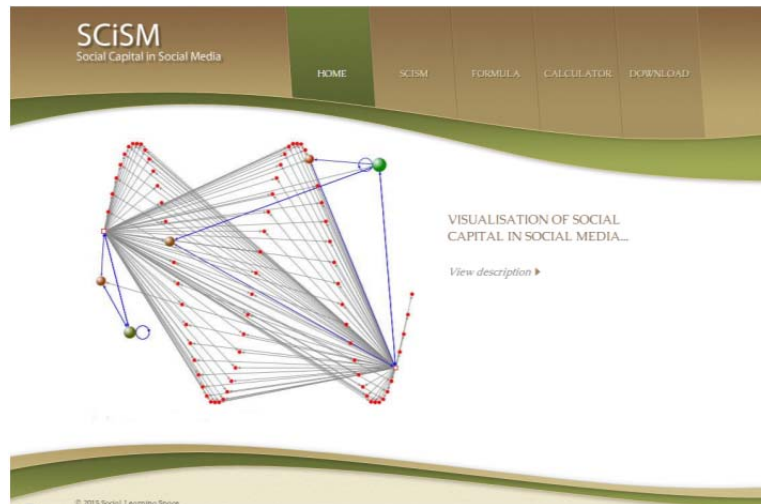


Figure 5.7 SCiSM web site

The web site contains a brief introduction to social capital theory and an outline of the SCiSM metric's formulae. In addition there is a 'SCiSM calculator' that enables other researchers to submit an XML file in order to view the results of the SCiSM metrics. The SCiSM software can also be downloaded from the web site.

5.5 Study 1: Validating the Proposed SCiSM Metrics

The aim of this thesis is to develop valid metrics for use in social media. Therefore the first step was to validate the proposed SCiSM metrics against existing valid measures taken from the literature i.e. Internet Social Capital Scale (ISCS) (Williams 2006). This type of validity is known as associative validity (Meneely 2012). The SCiSM metrics were validated using two independent data sets. The SCiSM metrics were also analysed for correlation to the Facebook Intensity Scale (FIS) (Ellison, Steinfield and Lampe 2006) as an experimental control. In other words the expectation is that the SCiSM metrics would correlate with ISCS but not with the FIS.

5.5.1 Aim of the study

The aim of this study is to validate the SCiSM metrics of social capital using two case study data sets, one from a Facebook group (data set 1) the second Facebook status messages (data set 2). Both data sets were analysed using ISCS scores as a baseline valid score. As an experimental control the SCiSM scores were also correlated with FIS scores, in order to eliminate the possibility the SCiSM was measuring Facebook use rather than social capital.

5.5.2 Methodology

In order to test for validity the SCiSM metrics social capital scores (for both data sets) were compared to ISCS. In addition, to act as a control, FIS scores were also tested for correlation to SCiSM scores.

The validation process using the Facebook group (data set 1) was carried out over twelve weeks by the eight volunteer participants. The eight participants were under-graduate computing students who were paid £10.00 in Amazon gift vouchers for taking part in this study. Originally there were ten volunteers but two did not complete the process. Each week of the twelve week period participants were asked to complete the ISCS and FIS surveys, the order was alternated. The participants were instructed to assess their answers to the surveys in respect of their use of the Facebook group in the preceding week.

The steps laid out in appendix D were used to prepare data set 1 for the SCiSM metrics calculations. It is worth highlighting the fact that users' experience of viewing messages was modelled in the data, as well as posting. Each week scores for the participants bonding and bridging social capital as related to use of the academic Facebook group was calculated using the SCiSM metrics and software. Scores were also calculated for weekly responses to ISCS and FIS.

The ISCS was adapted for use with a Facebook group see appendix E. The FIS survey see appendix F was created in order to measure intensity of use of Facebook.

The resultant scores were statistically analysed in terms of a participants SCiSM and ISCS as well as SCiSM and FIS. The statistical analysis was made up of correlations and linear regressions. The correlations were calculated using MS Excel with critical p-values of .05 relating to a score of 0.576 or above and .01 relating to a score of 0.708 or above (Kumar and Stevenson 1997) and (Stangroom 2014) and (Soper 2014). In addition a linear regression analysis was performed.

A second validation process was undertaken in order to determine whether or not the SCiSM metrics were generalisable to other social media contexts. The second validation used a data set based on Facebook status updates (data set 2). There were a total of 79 volunteer participants who responded to a request sent via email and Yammer. The request informed potential participants that they would be entered into a prize draw for a £20.00 Amazon gift voucher. The participants were asked to complete a survey comprising of the ISCS and questions concerning their recent use of the software. The survey is given in appendix G. The results for Data set 2 ISCS scores was calculated using the survey as was the SCiSM scores as explained in appendix H. The two data sets are given in appendix I. Data set 2 was statistically analysed in the same way as data set 1.

5.5.3 Results

The results explained in this section show a significant correlation and strong linear regression between the SCiSM metrics and the Internet Social Capital Scale (ISCS) (Williams 2006) using two data sets. It is argued that this indicates that the SCiSM metrics are valid metrics of social capital. However the correlation between the SCiSM scores and

the ISCS scores does not exclude that possibility that the SCiSM metrics may in fact be measuring another factor which happens to correlate with the ISCS. It is not practical to exclude all possible factors which might be correlating with ISCS however it is possible to determine if SCiSM metrics are measuring an aspect of using the Facebook software.

Therefore an examination of whether or not the SCiSM scores correlated with the Facebook Intensity Scale (FIS) (Ellison, Steinfield and Lampe 2006) was conducted. It was found that there was not a significant correlation between SCiSM and FIS.

5.5.3.1 Validation of proposed new metrics against ISCS Data Set 1

The research question addressed by the results in this section is:

Is there a statistical relationship between results obtained using a Facebook group (data set 1) for the ISCS and the scores produced using the SCiSM metrics?

The purpose of this research question is to determine whether SCiSM metrics can reasonably be argued as measuring social capital. The results of the correlation analysis between ISCS and the SCiSM metrics are given in the table 5.2 below.

Participant		ISCS	ISCSBond	ISCSBridge	Pvalues		
Participant 1	SCISM	0.84421	0.808372	0.7311489	01	01	01
	SCISMBond	0.832497	0.809566	0.7128885	01	01	01
	SCISMBridge	0.922188	0.739672	0.8924577	01	01	01
Participant 2	ISCS						
	SCISM	0.684372	0.594959	0.6740988	05	05	05
	SCISMBond	0.693254	0.604422	0.6809542	05	05	05
	SCISMBridge	0.569632	0.475393	0.5826405	No	No	05
Participant 3	ISCS						
	SC	0.869608	0.896919	0.7326441	01	01	01
	SCBond	0.865373	0.895491	0.7231394	01	01	01
	SCBridge	0.83553	0.780187	0.8686105	01	01	01
Participant 4	ISCS						
	SCISM	0.835431	0.844248	0.7745935	01	01	01
	SCISMBond	0.831645	0.842411	0.7684343	01	01	01
	SCISMBridge	0.851394	0.823012	0.8391324	01	01	01
Participant 5	ISCS						
	SCISM	0.776591	0.884429	0.5790092	01	01	05
	SCISMBond	0.757914	0.876293	0.5506601	01	01	No
	SCISMBridge	0.673613	0.489403	0.8072652	05	No	01
Participant 6	ISCS						
	SCISM	0.877578	0.896065	0.7999404	01	01	01
	SCISMBond	0.86826	0.88749	0.7906383	01	01	01
	SCISMBridge	0.899418	0.900611	0.8351318	01	01	01
Participant 7	ISCS						
	SCISM	0.857773	0.910603	0.7242457	01	01	01
	SCISMBond	0.799352	0.897057	0.6390314	01	01	05
	SCISMBridge	0.680867	0.406456	0.8090868	05	No	01
Participant 8	ISCS						
	SCISM	0.619876	0.651949	0.5352886	05	05	No
	SCISMBond	0.544007	0.608308	0.453957	No	05	No
	SCISMBridge	0.824191	0.710524	0.7801029	01	01	01

Table 5.2 Correlation between ISCS and SCiSM Metrics

The results of the analysis in table 5.2 shows the correlations between ISCS, column headings, and SCiSM scores in rows for each participant. The figures were arrived at by comparing the two scores for the 12 week period of the study. The key scores and p-values are highlighted. For example in reading the table the correlation between total ISCS and the total SCiSM metric for subject 1 is 0.84421 which is significant at .01 or in other words a 1 in 100 chance the results occurred by chance. In summary the null hypotheses can be rejected as all scores were correlated significantly at .05 or higher.

A linear regression analysis was also carried out. Scatter plot graphs were created for the SCiSM and ISCS scores, and the r^2 values calculated using MS Excel. Linear regression scores show the likelihood that one score predicts another score. The scatter plot graphs show an optimum line that scores should fall on if there is a perfect prediction.

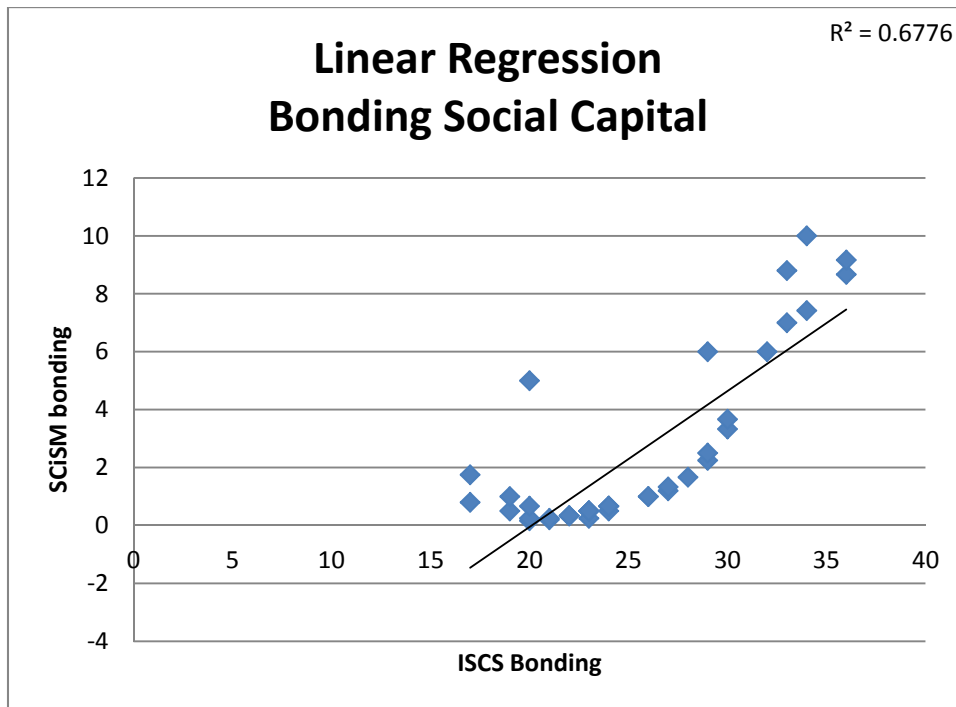


Figure 5.8 Linear Regression bonding ISCS scores and SCiSM Metrics for data set 1

The results for bonding social capital shown in figure 5.8 indicate a tendency for scores to fall near the optimum line. The data results in a relatively high score of $R^2 = 0.6776$.

However there is a fall away from the line for higher scores.

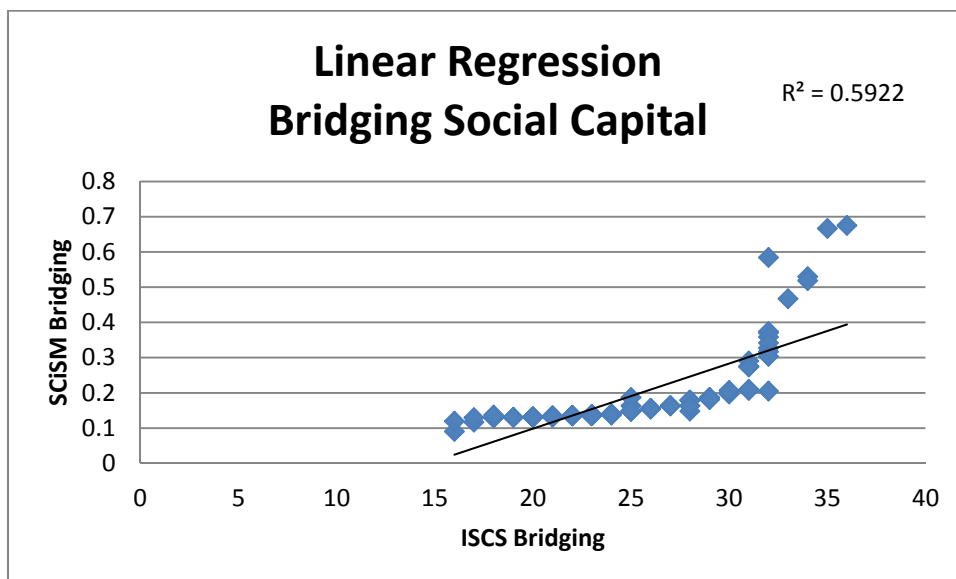


Figure 5.9 Linear Regression Bridging ISCS scores and SCiSM Metrics for data set 1

The results for bridging social capital shown in figure 5.9 indicate a tendency for scores to fall near the optimum line. The data results in another relatively high score of $R^2 = 0.5992$. However there is a clear fall away from the line for higher scores.

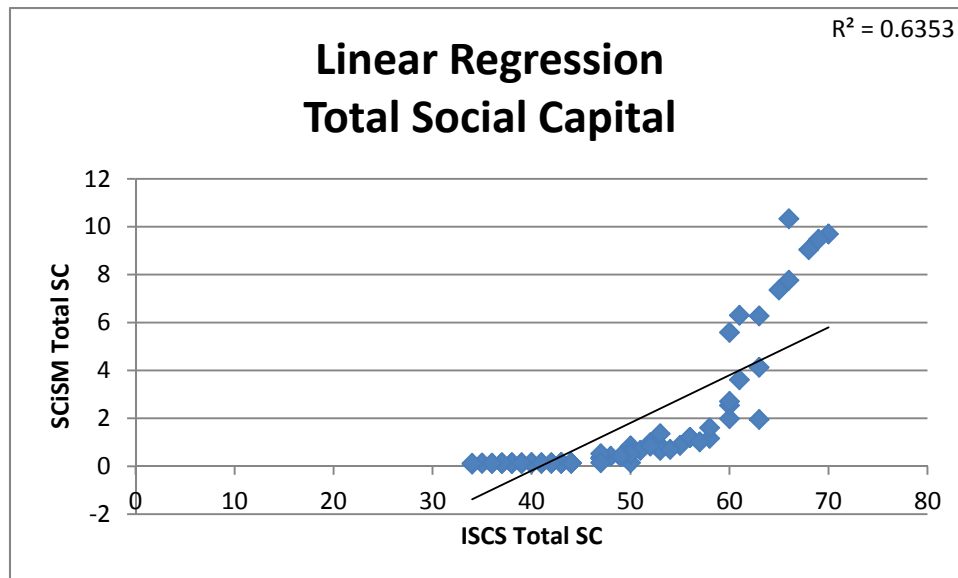


Figure 5.10 Linear Regression Total ISCS scores and SCiSM Metrics for data set 1

The results for total social capital shown in table 5.10 once again indicate a tendency for scores to fall around the optimum line. The data results in another relatively high score of $R^2 = 0.5992$. Again there is a clear fall away from the line for higher scores.

The results of the linear regression show that the **absolute value** of SCiSM scores is predictive of absolute value of ISCS scores up to a point. However there is a clear fall away in the predictive quality of the scores at the higher end.

Given this finding the participants' scores were then ranked for each score so that a table consisting of each participant's ranked SCiSM and ranked ISCS was produced.

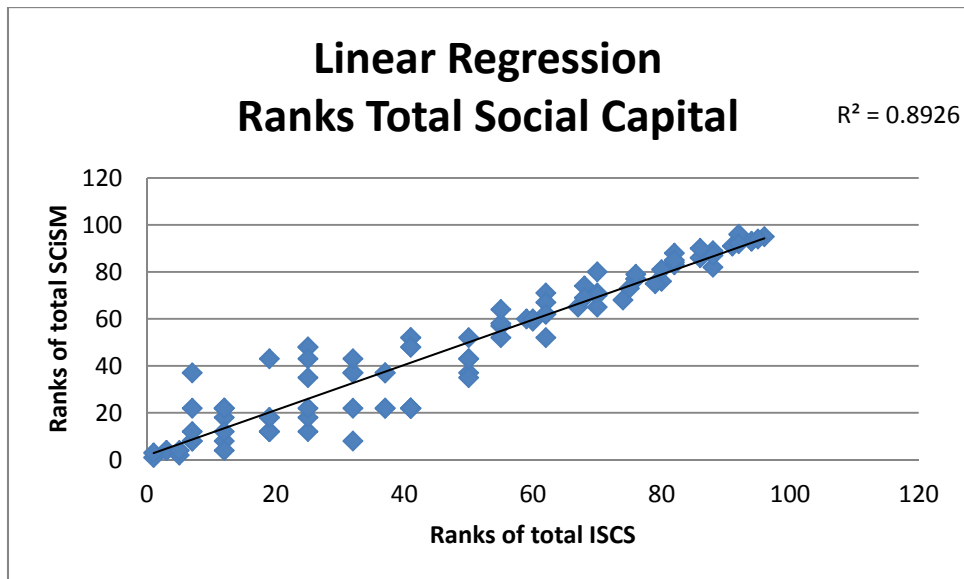


Figure 5.11 Linear Regression Total ISCS scores and SCiSM ranks for data set 1

The results for ranked total social capital shown in figure 5.11 show a stronger tendency for scores to fall around the optimum line. The data results in a high score of $R^2 = 0.8926$. The results suggest that the SCiSM metrics represent a better **relative value** score than for **absolute value** alone.

5.5.3.2 Validation of proposed new metrics against ISCS Data Set 2

The research question addressed by the results in this section is:

Is there a statistical relationship between results obtained using Facebook status updates (data set 2) for the ISCS and the scores produced using the SCiSM metrics?

The purpose of this research question is to determine whether SCiSM metrics can reasonably be argued as measuring social capital in more than one social media context.

The results of the correlation analysis between ISCS and the SCiSM metrics are given in the table 5.3 below.

Correlations	R	Significance
Bonding ISCS/SCiSM	0.687396	0.01
Bridging ISCS/SCiSM	0.72227	0.01
Total ISCS/SCiSM	0.782811	0.01

Table 5.3 Correlation between ISCS scores and SCiSM Metrics data set 2

The results are that there is a statistically significant $p=0.01$ correlation between ISCS and SCiSM scores for bonding, bridging and total social capital.

The results of a linear regression analysis on data set 2 followed the same pattern as data set 1.

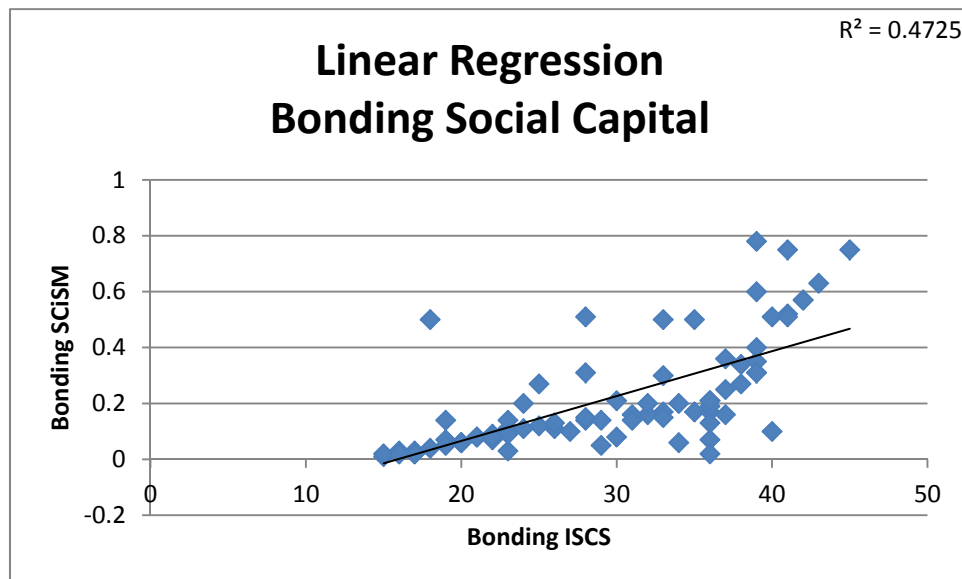


Figure 5.12 Correlation between bonding ISCS scores and SCiSM Metrics data set 2

The results for bonding social capital shown in figure 5.12 indicate a tendency for scores near the bottom range to fall near the optimum line. The data results in a score of $R^2 = 0.4725$.

However there is a fall away from the line for higher scores.

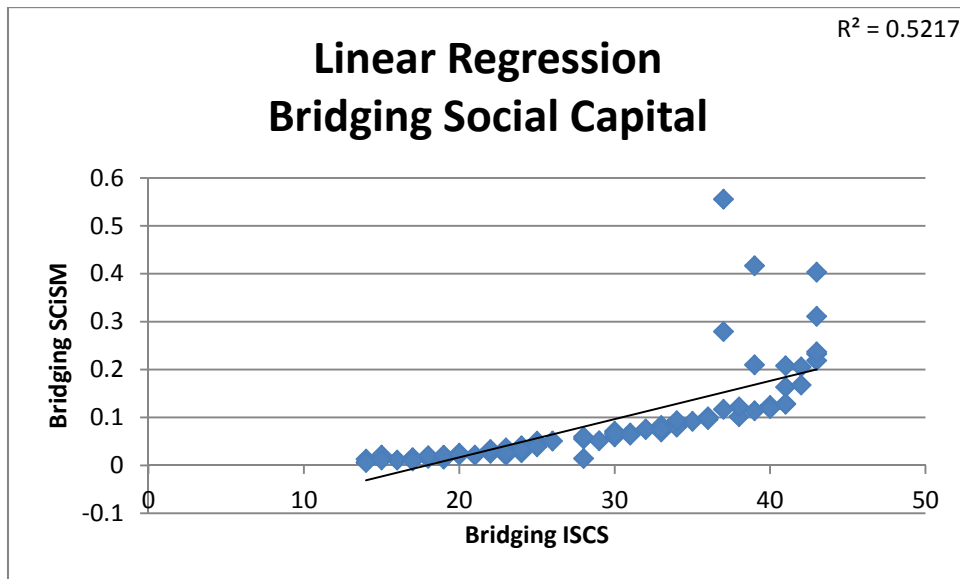


Figure 5.13 Correlation between bridging ISCS scores and SCiSM Metrics data set 2

The results for bridging social capital shown in figure 5.13 indicate a tendency for scores to fall near the optimum line. The data results in a relatively high score of $R^2 = 0.5217$.

However there is a fall away from the line for higher scores.

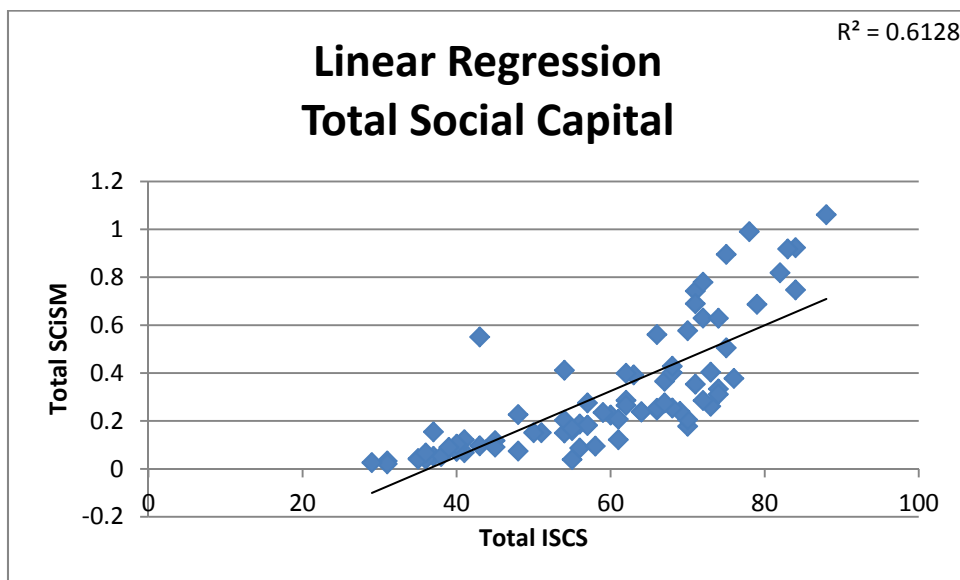


Figure 5.14 Correlation between Total ISCS scores and SCiSM Metrics data set 2

The results for total social capital shown in figure 5.14 indicate a tendency for scores to fall near the optimum line for the lower range of scores. The data results in a relatively high score of $R^2 = 0.6128$. However there is a fall away from the line for higher scores.

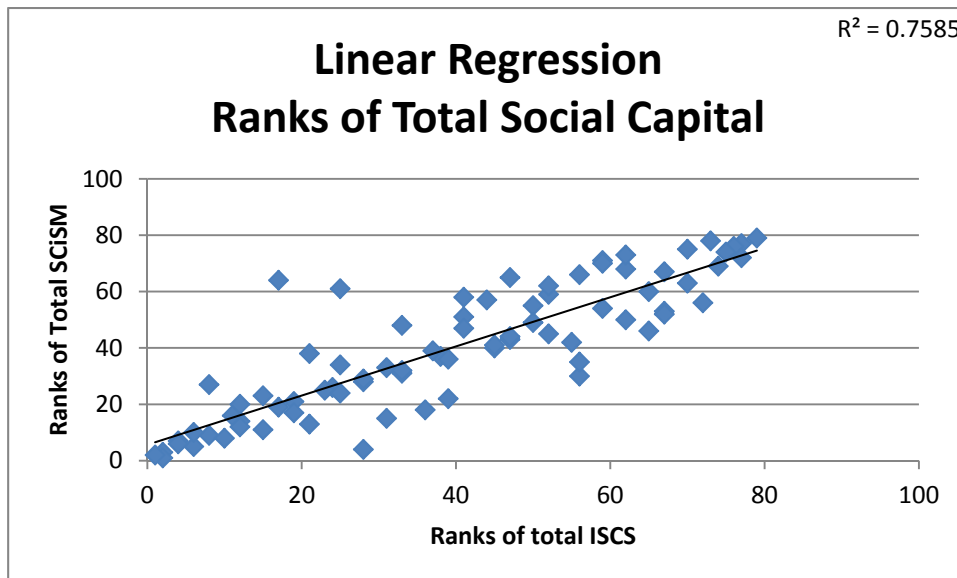


Figure 5.15 Correlation of ranks of total ISCS scores and SCiSM Metrics data set 2

The results for ranked total social capital shown in figure 5.15 show a stronger tendency for scores to fall around the optimum line. The data results in a high score of $R^2 = 0.7585$. The results lend further support for the assertion that the SCiSM metrics represent a better **relative value** score than for **absolute value** alone.

5.5.3.3 Analysis of proposed SCiSM metrics against FIS

The research question addressed in this section is:

Is there a correlation between scores for the Facebook Intensity Scale (FIS) (Ellison, Steinfield and Lampe 2006) and the scores produced using the SCiSM metrics on a Facebook group (data set 1)?

The purpose of this research question is to determine whether SCiSM metrics correlate with other Facebook metrics to act as an experimental control. The results of the correlation analysis between FIS and the SCiSM metrics are given in table 5.4 below. The results in the table are compiled from eight participants over a 12 week period.

Participant		FIS		Pvalues
Participant 1	SCISM	-0.18075		No
	SCISMBond	-0.19507		No
	SCISMBridge	-0.00299		No
Participant 2	SCISM	-0.27893		No
	SCISMBond	-0.28203		No
	SCISMBridge	-0.23807		No
Participant 3	SCISM	0.180733		No
	SCISMBond	0.178073		No
	SCISMBridge	0.223035		No
Participant 4	SCISM	-0.08881		No
	SCISMBond	-0.07951		No
	SCISMBridge	-0.25756		No
Participant 5	SCISM	0.12615		No
	SCISMBond	0.109462		No
	SCISMBridge	0.398177		No
Participant 6	SCISM	-0.08279		No
	SCISMBond	-0.08474		No
	SCISMBridge	-0.03145		No
Participant 7	SCISM	0.081455		No
	SCISMBond	0.056429		No
	SCISMBridge	0.19177		No
Participant 8	SCISM	0.410867		No
	SCISMBond	0.352707		No
	SCISMBridge	0.580328		05

Table 5.4 Correlation between FIS and SCiSM metrics

Table 5.4 shows the results of the analysis. In the table the column headings are for FIS scores and related p-values and total social capital using the SCiSM metrics in rows for each participant. The figures were arrived at by comparing the two scores for the 12 week period of the study. The p-values were tested for significance at .05 the required score is 0.576 and at 01 the required score is 0.708. There was only one significant correlation between the two sets of scores i.e. participant 8 SCiSM bridging score and FIS.

5.5.4 Discussion

The process of validation is essentially about constructing a study that lends support to an argument. In the case of this research the validation process involved a comparison of the SCiSM metrics against the existing and widely used Internet Social Capital Scale (ISCS) (Williams 2006) metric. The validation process involved comparing metric scores for two independent data sets. The scores were statistically analysed for correlations and regression analysis.

The first data set was derived from eight users of a Facebook group. There were two groups of data. The first group of data was obtained from the participants completing the ISCS survey for twelve weeks. The second group of data was interaction data from the Facebook group which was extracted and used to construct twelve weeks of SCiSM scores. The results of analysing the two groups of data indicate that there was a statistically significant correlation. The calculation of correlations involves ranking the data and arriving at a coefficient which is either significant or not. The correlation statistic is good at determining if there is a matching pattern in the data but it does not reveal anything about the actual pattern.

A linear regression analysis was carried out in order to reveal more about the pattern of similarity in the data. It was found that lower scores were closer to optimum than higher scores. Higher scores for ISCS were still associated with higher SCiSM scores but not on the optimum line. This finding suggests that if you were to substitute SCiSM for ISCS the absolute value of the scores would be less accurate than the relative score. This finding was further illustrated by ranking each data point within the two data groups and creating a scatter plot graph for a regression analysis. As expected this produced a far higher r score and a more consistent association of data points about the optimum line. This finding suggests that

SCiSM will provide an approximation for an absolute social capital score. However it provides a far more accurate relative score within a particular group of users.

The second data set was concerned with the use of Facebook status updates. Again there were two groups of data. The first group of data was obtained from a survey made up in part by the ISCS. The second group of data was derived from interactions using the software: the interactions where sending and receiving posts. The interaction data was used to calculate SCiSM scores. The two groups of data were analysed in the same way as the first data set i.e. correlations and regression analysis. The results of analysing the two groups of data indicate that there was a statistically significant correlation, as was the case with the first data set.

The results of the regression analysis on the second data were also similar to those of the first data set. Although, in the second data the variation from the optimum line for higher scores was less clear for bonding social capital than was the case for the first data set. However the pattern is more distinct for bridging social capital and total social capital. Overall, the second data set suggests that the SCiSM metrics are a valid measure particularly for relative scores.

The initial finding of validity leaves open another question, supposing the SCiSM metrics are just measuring user experience of using a Facebook group. If such a situation was the case we might expect that any measure of the user experience of the software would correlate with SCiSM. Therefore as an experimental control the SCiSM scores were examined against the Facebook Intensity Scale (FIS) (Ellison, Steinfield and Lampe 2006). The results of the analysis of correlations between the FIS vs SCiSM and were found to be not significant. This result lends weight to the idea that SCiSM measure social capital rather than Facebook use.

In conclusion the results indicate that the SCiSM metrics are a valid measure of social capital. In particular the SCiSM metrics provide a good relative measure of individual social capital. However it is worth noting that although it is arguable that the SCiSM metrics measure social

capital, it doesn't follow that they are the best measure of social capital in comparison to other Social Network Analysis (SNA) measures. Therefore the next section outlines a comparative evaluation of different measures against the SCiSM metrics.

5.6 Study 2: Evaluation of the SCiSM Metrics

The core aim of creating a measure of social capital in social media must, in addition to being valid, also take account of the metrics' relative accuracy in comparison to other SNA measures. This process of metric comparison was explained above as improvement validity i.e. that the new metrics should be an improvement on other measures.

5.6.1 Aim of the study

This study is concerned with examining how the SCiSM metrics differ from other measures such as: Degree Centrality, Betweenness Centrality, Clustering Coefficient and Eigenvector Centrality outlined above in chapter four. Put simply if other SNA measures of social capital are valid there is no need for a new measure. Therefore the research question addressed in this section is:

In a comparison between SCiSM scores and other SNA scores, which metrics validate most closely with ISCS?

The purpose of this research question is to determine whether the SCiSM metrics are as good as or better than existing SNA metrics.

5.6.2 Methodology

The SCiSM metrics were validated against ISCS scores, as indicated above, and found to have a significant positive correlation for all eight participants in the study. However in addition to SCiSM testing, SNA scores for degree, in-degree, out-degree, betweenness,

closeness, eigenvector centrality, and clustering co-efficient were compared to the ISCS scores. The performance of SCiSM's in terms of correlation to ISCS was tested against other SNA measures to determine whether it was in practice a more valid metric of social capital. The correlations were calculated using MS Excel with critical p-values of .05 relating to a score of 0.576 or above and .01 relating to a score of 0.708 or above (Kumar and Stevenson 1997) and (Stangroom 2014) and (Soper 2014). As was the case for the validation process for SCiSM, the SNA scores for the eight participants over 12 weeks were calculated using nodeXL and compared to the s participant's ISCS scores.

5.6.3 Results

The results of the correlations between the ISCS, ISCS bonding and ISCS bridging, the SNA measures and the SCiSM metrics for all participants are given in Appendix J. The next task was to broaden the comparison of the scores by summing the total number of significant correlations across all eight participants for all metrics.

	degree	in	out	Betweenne	Closeness	Eigen	Cluster	SCISM	SCISMBon	SCISMBridge
ISCS	6	6	7	3	0	3	7	8	7	7
ISCSBond	5	6	7	4	0	2	7	8	8	5
ISCSBridg	6	6	7	2	1	3	7	7	6	8

Table 5.5 Frequency of significant critical values for each score for all 8 participants

A score of eight indicates that the scores for all eight participants correlated for a specific metric. The results in table 5.5 show that the SCiSM metrics are significantly positively correlated for all eight participants, as indicated above during the validation process. The results also show relatively good performance for out-degree and clustering co-efficient scores with seven participants out of eight, correlating with the ISCS scores. However these scores don't correlate as well as the SCiSM metrics do for bonding and bridging social capital. These results are discussed in the next section.

5.6.4 Discussion

The results shown in table 5.5 show that the SCiSM scores were significantly correlated with the ISCS for all eight participants. The results concerning degree centrality correlations to ISCS scores show a relatively good performance. These scores were, in terms of number of significant correlations; ISCS with six, ISCS bonding with five and ISCS bridging with six. In the instances where there was not a significant correlations, for example participant 8, there were relatively few posting. This suggests that there are limitations to the usefulness of degree centrality, which is a network traffic score, for situations where there is light traffic. The results concerning degree centrality scores are particularly interesting due to the fact that this measure has been used extensively in the literature. The in-degree and out-degree scores for each of the ISCS scores were 6's and 7's respectively. These relatively high numbers of correlations were the reason these scores were used in the SCiSM metrics' design. These findings suggest that the SCiSM metrics are a more consistent measure of bonding, bridging and total social capital.

The betweenness centrality measure is the effect on the graph if a node is removed.

Betweenness centrality is calculated by determining the shortest path between the nodes and then for each pair of nodes. The number of times that betweenness centrality is significantly correlated as shown in table 5.5 is ISCS three, ISCS bonding four and ISCS bridging two.

This result is not surprising as it seems unlikely that a participant's experience of using the software encompasses the idea of their relative connectedness to other participants. In other words if you were to post, comment or read messages in a Facebook group there are several things you might have an awareness of, for example the amount of times another person posts to the group. However it seems unlikely that you would have a clear idea about the effect on

the group if particular people were removed. Furthermore the number of correlations between betweenness centrality score and ISCS score is quiet low.

The closeness centrality score is based upon the idea of the shortest path across the graph, and is a measure of connectedness. There was only one instance where closeness centrality score was significantly correlated to an ISCS score and that was for bridging ISCS. As was the case for betweenness centrality this is likely to be because the participant's experience of using the software would not encompass the idea of their connectivity to other nodes on the graph.

The eigenvector centrality score is essentially a measure of connectedness to popular nodes. The number of times that eigenvector centrality is significantly correlated as shown in table 5.5 is ISCS three, ISCS bonding two and ISCS bridging three. This result suggests that interactions with 'popular' members of a group do not correlate with an individual's perception of their own social capital.

The clustering coefficient score is a clique score where connectedness to other nodes is assessed in triads. The process of arriving at a score involves grouping connected nodes into three and determining if the triad has three sides, i.e. each node connected to the other two. The results show that there were significant correlations for ISCS total with seven, ISCS bonding with seven and ISCS bridging with seven. These results suggest that there are limitations to the usefulness of clustering coefficient scores. The clustering coefficient score for participant eight failed to correlate with total ISCS and ISCS bridging scores. Participant eight made a relatively few contributions. The SCiSM bonding metric is basically a refinement of the clustering coefficient score, designed to detect bonding social capital even in instances where there is relatively little network traffic.

In conclusion the evaluation between SNA scores and SCiSM scores provides an interesting set of results. The scores for out-degree and clustering coefficient were all relatively high, whilst the scores for betweenness centrality, closeness centrality and eigenvector centrality were much lower. It has been explained that the SCiSM metric for bridging social capital is essentially a refinement of degree centrality. Furthermore the SCiSM metric for bonding social capital is essentially a refinement on clustering coefficient. The key point is that the **SCiSM metrics** correlated in all instances and therefore **represent an improvement on the other SNA measures**.

5.7 Study 3 Using SCiSM to analyse social media: Facebook Group

The purpose of conducting this study was to demonstrate what the SCiSM metrics can reveal about users' experience of social media. In particular this study concerning a Facebook Group illustrates how SCiSM can be used to analyse the social capital of individuals in a user group. The results from the study are related to key findings from the literature review.

5.7.1 Aim of study

The aim of this study is to examine if there was a change in the social capital of users of a Facebook group between the first and second half of a twelve week period.

5.7.2 Methodology

Data set 1 which was used in the validation of the SCiSM metrics was also used in this study. From the data for each participant the three SCiSM scores were summed for the first and last six weeks of a twelve week period. The resultant scores for each participant were then analysed making use of a t-test.

5.7.3 Results

The data was imported into MX Excel and t-tests were performed. The paired t-test selected was a 'paired two sample for means' as there were two samples for each participant. A paired t-test examines the variation of values within each sample, and produces a single number known as a t-value. The next step is to compute a p-value which is the probability that two samples from the same population would produce a t-value such as the one found. Therefore, a t-test measures how different two samples are (the t-value) and tells you how likely it is that such a difference would appear in two samples from the same population (the p-value). A one tailed test was used because it is expected that scores would go up. The results of the the t-tests are given below.

t-Test: Paired Two Sample for Means			
	<i>last six weeks</i>	<i>first six weeks</i>	
Mean	14.29027868	4.204527414	mean for variable 1 is higher
Variance	227.5681877	10.0022556	variance for variable 1 is higher
Observations	8	8	
Pearson Correlation	0.740414749		
Hypothesized Mean	0		
df	7		
t Stat	2.207994477		t-Test statistic > critical value
P(T<=t) one-tail	0.0314917		
t Critical one-tail	1.894578604		Reject the null hypotheses
P(T<=t) two-tail	0.0629834		
t Critical two-tail	2.364624251		

Figure 5.16 Results of a t-test on total SCiSM scores for first vs second six weeks

Figure 5.16 shows the eight participant total SCiSM scores for two time periods i.e. the first six weeks and last six weeks produced by the SCiSM software. Therefore there are eight pairs of scores, the first of the pairs is for weeks one to six, and the second of the pairs are for week's seven to twelve. Figure 5.16 show the results of examining the two time periods in terms of total SCiSM scores using a t-test. The results show that the mean total SCiSM

scores are significantly higher, in a one tailed test at $p=.05$, in the last six weeks than the first six weeks.

t-Test: Paired Two Sample for Means			
	<i>last six weeks</i>	<i>first six weeks</i>	
Mean	13.0828869	3.083333333	mean for variable 1 is higher
Variance	218.8427057	9.404761905	variance for variable 1 is higher
Observations	8	8	
Pearson Correlat	0.705646431		
Hypothesized M	0		
df	7		
t Stat	2.207042925		t-Test statistic > critical value
P(T<=t) one-tail	0.03153596		
t Critical one-tail	1.894578604		Reject the null hypotheses
P(T<=t) two-tail	0.063071921		
t Critical two-tail	2.364624251		

Figure 5.17 Results of a t-test on SCiSM Bonding scores for first vs second six weeks

Figure 5.17 shows the results of examining the two time periods bonding SCiSM scores using a t-test. The mean bonding SCiSM scores are significantly higher, in a one tailed test at $p=.05$, in the last six weeks than the first six weeks.

t-Test: Paired Two Sample for Means			
	<i>last six weeks</i>	<i>first six weeks</i>	
Mean	1.207391773	1.12119408	mean for variable 1 is higher
Variance	0.148688053	0.044691671	variance for variable 1 is higher
Observations	8	8	
Pearson Correlation	0.436968999		
Hypothesized Mean	0		
df	7		
t Stat	0.697612918		t-Test statistic NOT > critical value
P(T<=t) one-tail	0.253960507		
t Critical one-tail	1.894578604		Can not reject the null hypotheses
P(T<=t) two-tail	0.507921014		
t Critical two-tail	2.364624251		

Figure 5.18 Results of a t-test on SCiSM Bridging scores for first vs second six weeks

Figure 5.18 show the results of examining the two time periods bridging SCiSM scores using a t-test. The mean bridging SCiSM scores are not significantly higher, in a one tailed test at $p=.05$, in the last six weeks than the first six weeks.

In summary the results of the comparisons between the first half and the second half of a twelve week period using a t-test were as follows:

- Total SCiSM increased significantly in the second half
- Bonding SCiSM increased significantly in the second half
- Bridging SCiSM increased but not significantly in the second half

These results are discussed in the next session.

5.7.4 Discussion

The results of this study show that there was a significant increase in total and bonding SCiSM scores, from the first six weeks to the second six weeks of a twelve week time period. There was also an increase in the bridging SCiSM score but this was not statistically significant. In the literature review it was stated that there was a key finding that Steinfield et al (2009) had found i.e. that bonding social capital increased over time amongst users of social media but bridging did not. Therefore the result from this study supports the Steinfield et al (2009) assertion.

In the case of this study it should be borne in mind that the use of the Facebook group was within the setting of the first semester of an academic calendar. The results of this study may simply be reflecting that participants found the Facebook group performed better as a means of accessing bonding social capital after several weeks experience in using the software. On the other hand, the results suggest that the Facebook group software is performing well as a means of improving bonding social capital between users, but less so in terms of improving bridging social capital.

The result of this study and Steinfield et al (2009) has particular significance if the intention of using the software was to act as an icebreaker allowing user to create relationships within a group. However if the intention is to improve information flow, as is suggestive of the fact bridging social capital did not increase significantly, then it would appear the further steps are required. It is also possible that both studies are detecting a usage pattern that may appear in many other contexts and as such is worthy of further study.

The value of the SCiSM metrics is twofold. Firstly as a contribution to the body of knowledge, the metrics are able to produce comparable results shedding light on other research findings. Secondly the SCiSM metrics can be of value to user groups who wish to have a greater understanding of the social capital that social media can support.

5.8 Study 4 Using SCiSM to analyse social media Facebook status updates

The purpose of conducting this study was to further demonstrate what the SCiSM metrics can reveal about users' experience of social media. In particular this study concerning Facebook friends illustrates how SCiSM can be used to analyse key features in social media. The results from the study are related to key findings from the literature review.

5.8.1 Aim of the study

The aim of this study is to examine if there is a relationship between the social capital of users of Facebook, their number of Facebook friends and their posting of a Facebook status update.

5.8.2 Methodology

Data set 2 used in the validation of the SCiSM metrics was used in this study. As explained above the data was extracted from a survey n=79, which contained the ISCS questions. The

survey also contained questions on the user number of Facebook friends as well as their recent posting to Facebook.

5.8.3 Results

The number of Facebook friends and SCiSM scores were statistically tested for correlation. The results were arrived at by comparing the number of Facebook friends in turn with bonding, bridging and total SCiSM. The p-values were tested for significance at .05 the required score is 0.2214 for n=79. The results are shown in table 5.7.

		Significance
Friends & Total	0.080883	No
Friends & Bonding	0.249347	0.05
Friends & Bridging	-0.27856	0.05

Table 5.6 Number of Facebook friends vs SCiSM scores correlations

In results in table 5.6 shows that the correlation between number of Facebook friends and total SCiSM is 0.080883 is not significant at .05, or in other words a 5 in 100 chance the results occurred by chance.

Table 5.6 also shows that the correlation between number of Facebook friends and bonding SCiSM is 0.249347 is significant at .05.

Table 5.6 also shows that the correlation between number of Facebook friends and bridging SCiSM is -0.27856, a negative correlation, is significant at .05.

The number of Facebook status update posts and SCiSM scores were statistically tested for correlation. The results were arrived at by comparing the number of Facebook status posts in turn with bonding, bridging and total SCiSM. The p-values were tested for significance at .01 the required score is 0.288 for n=79. The results are shown in table 5.8.

		Significance
Posting & Total	0.440925	0.01
Posting & Bonding	0.442152	0.01
Posting & Bridging	0.291289	0.01

Table 5.7 Number of Facebook posts vs SCiSM scores correlations

In results in table 5.7 show that the correlations between number of Facebook status post and total, bonding and bridging SCiSM are significant at .01.

5.8.4 Discussion

The results indicate that there is a positive, statistically significant, correlation between number of friends and bonding social capital (as measured by SCiSM). In addition there is a negative correlation between number of friends and bridging social capital. This result indicates that the more Facebook friends a user has the less value is being derived from the broader information network as represented by bridging social capital. Therefore it is suggested that this result indicates that there is a **dilution effect** in the bridging social capital of Facebook users.

There was not a statistically significant correlation between total social capital (measured by SCiSM) and number of Facebook friends. Given that bonding social capital had a positive correlation in one direct and bridging social capital a negative one in the other direction, it is not surprising that total social capital has arrived at close to zero.

The results show that there is a correlation between posting interactions with Facebook and bonding, bridging and total social capital (as measured by SCiSM). In assessing this result it is important to bear in mind that correlation do not imply that there must be a causal relationship between two data points. However correlations do not exclude the possibility

that there is a causal relationship. Therefore there is a possibility that increasing Facebook posting can increase the users' experience of social capital.

The findings from this study can shed light on the key findings from the literature outlined above. For example two studies in the literature suggested that social capital can be measured by number of friends (Zhao 2006; Pfeil, Arjan and Zaphiris 2009). Whilst on the other hand Ellison, Steinfield and Lampe (2011) suggested that this was true only up to a point. The results from this study lend support to the Ellison, Steinfield and Lampe (2011) assertion that number of Facebook Friends cannot serve as a proxy for social capital.

In addition the literature review revealed that in a recent study by Ellison et al (2014) it was found that social capital was not generated by data such as friends but rather by interactions such as postings (Ellison et al. 2014). This study had a similar result in terms of posts.

It is suggested that the SCiSM metrics can make a valuable contribution to social capital theory. For instance the contribution to the body of knowledge, the metrics are able to produce e.g. support for a **dilution effect** in bridging social capital for users with a large number of friends. Secondly the SCiSM metrics can be of value to the research community by providing a tool to facilitate greater understanding of the social capital users can obtain from social media web sites such as Facebook.

5.9 Chapter Summary

The chapter began with an explanation of the methodology used for validating the social capital metrics which was based upon Meneely (2012). Original metrics called Social Capital in Social Media (SCiSM) was put forward and explained as being divided into three distinct elements relating to bonding, bridging and total social capital. This chapter also contained a

description of how the SCiSM metrics were implemented in software using an extensible MVC architecture.

The first study in this chapter outlined the successful validation process for the SCiSM metrics against ISCS (Williams 2006) using two data sets. The first study also contained an experimental control test which involved testing to see if SCiSM correlated against the Facebook Intensity Scale (FIS) (Ellison, Steinfield and Lampe 2006), which it did not. The FIS result lends support to the argument that SCiSM measures social capital and not simply usage of Facebook.

In the second study the SCiSM metrics were evaluated against other SNA measures such as degree centrality. The evaluation took the form of correlations against the ISCS (Williams 2006). It was found that SCiSM had the best performance as compared to the other SNA measures, particularly when broken down into bonding and bridging social capital.

The SCiSM metrics are refinements of existing SNA measures and they have an underlying logic that suggests that they are measuring social capital. In other words for bonding social capital a measure based upon clustering coefficient has both logic and validity. This is because clustering coefficient scores are measuring cliqueness or in other words the tightness of bonds between users. The same can be said for bridging social capital as it is a measure based upon in-degree. In the case of bridging SCiSM the underlying logic is that in-degree is a measurement of a user reception of the flow of information. It is suggested that SCiSM metrics measure the social capital that users experience when using social media software.

The chapter also contained the results of two studies which demonstrated how the SCiSM metrics can be used to examine the experience of social media. The first study, using a Facebook group, found that total SCiSM and bonding SCiSM increased significantly in the

two consecutive time periods which were examined. Furthermore it was found that bridging SCiSM scores increased but not at a statistically significant level. The second study, using Facebook status updates, found amongst other things that number of Facebook friends cannot be used as a proxy score for social capital.

Chapter 6 Conclusions

6.1 Introduction

The overall goal of this research has been to develop automated metrics of social capital for use with social media software. This final chapter contains an explanation of the main conclusions, a summary of the thesis and a discussion of future work.

6.2 Thesis Conclusions

The aim of this research was to **create, validate and evaluate social capital metrics in order to analyse the social interaction performance of social media**. In responding to this aim the conclusions were:

- The literature review found that taken together Social Capital theory and Social Network Analysis (SNA) techniques could address the aim. In addition the concepts of bonding and bridging social capital (related to friends and acquaintances) were relevant.
- A preliminary analysis making use of a Facebook group found that degree centrality and clustering coefficient metrics might be useful in an analysis of Social Media.
- It was proposed that the Social Capital in Social Media (SCiSM) metrics could be calculate as:
 - Bonding Social Capital = out-degree multiplied by clustering coefficient
 - Bridging Social Capital = actual in-degree divided by potential out-degree
 - Social Capital = Bonding Social Capital + Bridging Social Capital
- The SCiSM metrics were successfully validated against two independent data sets using the Internet Social Capital Scale by means of correlations and regression analysis.
- The SCiSM metrics were evaluated against other SNA measures. The results were that SCiSM was superior to betweenness, closeness and eigenvector centrality and an improvement on degree centrality and clustering coefficient.

- The SCiSM metrics were demonstrated in practice using the two independent data sets in order to show their utility. It was found that the metrics showed
 - In a Facebook group there was an increase over time in bonding social capital but not bridging social capital
 - In the case of Facebook status updates there was a positive correlation between bonding social capital and number of Facebook friends but a negative correlation between bridging social capital and Facebook friends. The results for bridging social capital and Facebook friends represents a dilution effect suggesting that there are limits to the usefulness of the ‘friends’ facility. It was also found that posting was positively correlated with both bonding and bridging social capital.

The conclusions are further elucidated in the following thesis summary.

6.3 Thesis Summary

The literature review was comprised of three key components which were: ‘social capital theory’, ‘social capital in social media’ and the ‘measurement of social capital in social media’. In chapter two the literature review suggested that social capital theory could provide a framework for understanding the structure that exists in a social network. It was explained that there are a number of studies which have advanced the understanding of how computer based communication can support social capital. It was also suggested that Putnam (1995, 2000) was a particularly influential contributor to the literature having introduced the ideas of bonding and bridging social capital, which relate to friends and acquaintances respectively. It was suggested that the concepts of bonding and bridging social capital were relevant to this study of social media metrics.

The literature review continued in the third chapter by focusing on how social capital has been measured in social media research. The analysis in the literature review sought to determine: precedence, guidance, omission and key findings. It was argued that there were

ample precedents for the use of social capital theory as applied to a social media context. In addition it was explained that the literature provided guidance for this research. For instance it was found that social capital is often measured by means of surveys which includes the influential Internet Social Capital Scale (ISCS) (Williams 2006) and the Facebook Intensity Scale (FIS) (Ellison, Steinfield and Lampe 2006). In addition social capital is also frequently measured by degree centrality which it was explained is a method from Social Network Analysis (SNA). Given that the intended result of this research is the development of metrics and related software it was concluded that SNA was worthy of further consideration. It was also explained that there is an omission, or gap, in the literature in so far as there is a lack of valid SNA based metrics for measuring social capital. Furthermore given that the intention of this research is to create metrics for social capital it was thought that describing key findings from the literature would provide a context to demonstrate the utility of the SCiSM metrics. From the literature the findings highlighted included the development of bonding social capital over time. In addition the findings also included the relationship between number of Facebook friends and social capital. These findings from the literature review were examined in the light of results found using the SCiSM metrics explained in chapter five. The conclusions from the literature review included the argument that social capital theory and SNA were appropriate for use in the analysis of social media.

The second research aim namely “Preliminary analysis of social capital in social media using a case study” was addressed in the fourth chapter. The chapter contained the preliminary analysis of a Facebook group used in an academic setting. It was found that degree centrality provided a means of identifying the structural relationships in the social graph. In addition clustering coefficient scores suggested that students preferred to communicate with users who were identifiable as cliques of friends. In the light of these findings it was suggested a student Facebook group was a suitable test bed for the development of social metrics and

related software. The fact that the software has a dual purpose of providing information and access to friends suggests that it might map successfully to the concepts of bridging and bonding social capital proposed by Putnam (1995, 2000). The fact that similar patterns emerged between the three Facebook groups also suggests that results from one group may be generalisable to other groups.

The final research aim “Develop the metrics and related software to measure social capital for use with social media data” was addressed in the fifth chapter. It was explained that the process of developing the social capital metrics would involve: proposition, implementation, validation, evaluation and demonstration.

The metrics, which were referred to as Social Capital in Social Media (SCiSM), were outlined as formula for bonding, bridging and total social capital. Chapter five also contained a description of how the SCiSM metrics were implemented in software in order to make available the metrics to other researchers.

The process of arriving at valid metrics for social capital was crucial in order to ensure the metrics’ credibility and utility. Therefore a methodology for validating the metrics, based upon Meneely (2012) was utilised. The notion of validity is fairly fluid; in a sense a metric is valid if it can be reasonably argued to be so. However, it is suggested, that by utilising multiple criteria associated with validity and implementing those criteria in a systematic methodology the argument that a metric is valid can be considerably strengthened.

It was essential to determine that SCiSM measures a user’s experience of social capital and not merely some other aspect of social interaction. Therefore the metrics were validated against the Internet Social Capital Scale (ISCS) (Williams 2006). The notion of associative validity was utilised in examining the SCiSM metrics against the ISCS. There were two data

sets used in the validation process. The first data set was comprised of eight participants in a weekly test running for twelve weeks. Each week the participants completed the ISCS survey and a calculation of SCiSM scores was taken from the Facebook group the participants were using. The statistical tests used in the validation process were correlations and linear regression analysis. It was found that there was a significant correlation between the SCiSM scores and ISCS scores. The linear regression analysis showed that SCiSM was more effective as a relative score than an absolute score of social capital. The validation was repeated on a second data set, which was based upon Facebook status updates from seventy nine participants. The results from the second data set were the same as for the first.

In order to further strengthen the argument that social capital was being measured and not merely Facebook use, SCiSM was tested against the Facebook Intensity Scale (FIS). It was found that there was not a correlation between the two.

The fifth chapter also contained a study examining the performance of the proposed new SCiSM metrics against other Social Network Analysis (SNA) based measures. The performance was based upon the number of significant correlations between each of the measures and the ISCS. The results showed that SCiSM was significantly correlated with the ISCS for all eight participants. Furthermore the metrics were broken down into its component parts of bonding and bridging SCiSM and these were also significantly correlated with the ISCS for all eight participants. None of the other SNA measures performed at the same level as SCiSM.

The third and fourth studies outlined in chapter five demonstrated the use of SCiSM in measuring social capital using the two data sets used in the validation process. The intention was to illustrate how SCiSM metrics can be used to contribute to the existing literature on measuring social capital in social media. The literature in question was highlighted as key

findings in the literature review. The results of the SCiSM metrics study of Facebook groups indicated that bonding social capital increased over time in line with findings from Steinfield et al (2009). In addition there was a second study concerning the relationship between numbers of Facebook friends and social capital. The study in this research found that there was a positive correlation between bonding social capital and numbers of Facebook Friends, whilst there was a negative correlation between bridging social capital and numbers of Facebook Friends. This result was described as a dilution effect for bridging social capital and shed further light on key findings. For example two studies in the literature suggested that social capital can be measured by number of friends (Zhao 2006; Pfeil, Arjan and Zaphiris 2009). Whilst on the other hand Ellison, Steinfield and Lampe (2011) suggested that this was true only up to a point. Therefore the results from this study lend support to the idea that number of Facebook Friends cannot serve as a proxy for social capital.

In summary, the overall aim of this research has been to develop, and implement in software, a valid metrics to analyse the performance of social media. In order to fulfil this aim an innovative process of using social capital theory to inform the logic behind the metrics was used. The aim was achieved by the creation of original a validated metrics that enable the automated measurement of social capital in social media software. The thesis also contributed to the existing literature by demonstrating results obtained from the metrics.

6.4 Future Research

In terms of future development of the SCiSM metrics, they have been shown to be a valid measure of social capital. However linear regression analysis suggested that the SCiSM metrics were less reliable in instances of high scores. It would therefore be useful to analyse the threshold where the metrics became less reliable. For instance the calculation for the

metrics could be adjusted to take account of threshold set for the top 5% of the sample population and then introduce a dampening factor to those scores.

There are potential research questions concerning the utility of individual social media platforms as a means of allowing people to develop social capital over time. For instance research could be conducted in educational settings. In this study it has been demonstrated that a student user groups could gain insight into the development of social capital over the course of a period of study. In the future measurements of social capital could be taken before and after specific pedagogic strategies in order to assess their effectiveness. Furthermore the SCiSM metrics could be used to investigate relationship maintenance over time in order to examine the findings from Ellison et al (2014).

There are also potential research questions relevant to software developers concerning the various social media platforms. For example the SCiSM metrics could also be used to analyse whether or not twitter supports the development of social capital. In addition to these general questions about social media platforms there are possibilities in examining how the various functions and user interface elements in social media affect users' social capital.

It may also prove useful to deploy the SCiSM metrics in a business context by examining social capital in organisations. For instance the metrics could be used to analyse the development of bridging social capital as a means of improving information flows within an organisation. Furthermore the metrics could be used to examine the findings from studies such as Sun & Shang (2014) which suggested the increases to social capital improved productivity.

There are also potential applications for the SCiSM metrics in the field of criminology. Whilst it is unlikely that a social capital score alone could help to detect criminals, the

SCiSM metrics could help to provide intelligence of criminal gangs and networks.

Furthermore it may be possible to use SCiSM to find influential people and to map the spread of ideas which could be valuable in counter-terrorism.

Overall these directions for future research suggest the SCiSM metrics may provide a useful tool for future researchers.

References

- Adler, P and Kwon, S 2002, Social capital: prospects for a new concept, *Academy of Management Review*, 27, (1).
- Alexa Internet 2015, Top sites, *Alexa the web information company*, viewed 1 June, 2013, <<http://www.alexa.com/topsites>>.
- Allport, G 1954, *The nature of prejudice*, Addison-Wesley, Cambridge MA.
- Antoci, A, Sabatini, F and Sodini, M 2012, See you on Facebook! A framework for analyzing the role of computer-mediated interaction in the evolution of social capital, *The Journal of Socio-Economics*, 41, (5), Elsevier Inc.
- Appel, L, Dadlani, P, Dwyer, M, Hampton, K, Kitzie, V, Matni, Z a, Moore, P and Teodoro, R 2014, Testing the validity of social capital measures in the study of information and communication technologies, *Information, Communication & Society*, 17, (4), Taylor & Francis.
- Arregle, J, Hitt, MA and Sirmon, DG 2007, The Development of Organizational Social Capital : Attributes of Family Firms, *Journal of management studies*, (January).
- Arvidsson, A 2011, General sentiment - how value and affect converge in the information economy, *Social Science Research Network*.
- Austin, D 2013, How Google Finds Your Needle in the Web's Haystack, *American Mathematical society*, viewed <<http://www.ams.org/samplings/feature-column/fcarc-pagerank>>.
- Baker, W 1990, Market networks and corporate behavior, *American Journal of Sociology*, 96.
- Baker, W and Faulkner, R 2009, Social capital, double embeddedness, and mechanisms for stability and change, *American Behavioral Scientist*, 52, (11).
- Baym, NK 2013, First Monday , Volume 18 , Number 10 - 7 October 2013, *First Monday*, 18, (10).
- Beauchamp, T and Childress, J 2001, *Principles of Biomedical Ethics*, 5th ed, Oxford University Press.
- Bermejo, F 2007, *The Internet Audience*, Peter Lang, New York, USA.
- Best, SJ and Krueger, BS 2006, Online Interactions and Social Capital: Distinguishing Between New and Existing Ties, *Social Science Computer Review*, 24, (4).
- Borgatti, S, Everett, MG and Freeman, LC 1999, *UCINET 5 for Windows*, Analytic Technologies, MA.

- Borgatti, SP, Jones, C and Everett, MG 1998, Network measures of social capital, *Connections*, 21, (2).
- Bourdieu, P 1980, Le capital social: notes provisoires, *Actes Rech. Sci. Soc.*
- Bourdieu, P 1996, On the family as a realised category, *Theory, Culture and Society*, (13).
- Bourdieu, P 1985, The forms of capital, *Handbook of theory and research*, Greenwood, New York, USA.
- Bourdieu, P and Wacquant, L 1992, An invitation to reflexive sociology, University of Chicago, Chicago.
- Brandes, U 2001, A Faster Algorithm for Betweenness Centrality, *Journal of Mathematical Sociology*, 169, (5).
- Briand, L, El Emam, K and Morasca, S 1995, Theoretical and empirical validation of software product measures, *International Software Engineering Research Network*.
- Burke, M, Kraut, R and Marlow, C 2011, Social Capital on Facebook : Differentiating Uses and Users, *CHI 2011*.
- Burt, R 1992, Structural Holes, The Social Structure of Competition, Harvard Univ. Press, Cambridge, USA.
- Butler, T 2012, Model-View-Confusion part 1: The View gets its own data from the Model, viewed <<https://r.je/views-are-not-templates.html>>.
- Catanese, SA, Meo, P De, Ferrara, E, Fiumara, G and May, SI 2011, Crawling Facebook for Social Network Analysis Purposes.
- Checkfacebook 2015, socialbakers.com, viewed 1 June, 2013, <<http://www.socialbakers.com/statistics/facebook/>>.
- Cohen, D and Prusak, L 2001, In good company: How social capital makes organizations work, *The Economist*, 154.
- Coleman, J 1990, Foundations of Social Theory, Belknap Press of Harvard University Press, Cambridge, USA.
- Coleman, J 1988, Social Capital in the creation of Human Capital.
- Craven, M 2013, Discussions on Social Network Analysis.
- Dakhli, M and DeClercq, D 2003, Human Capital, Social Capital and Innovation: A Multi-Country Study, *International Business*, 21, (1).
- Daniel, B, Schwier, R and McCalla, G 2003, Social Capital in virtual learning communities and distributes communities of practice, *Canadian Journal of Learning Technology*, 29, (3).

- Davis, J and Leinhardt, S 1972, The structure of positive relations in small groups, in J Berger, M Zilditch & B Anderson (eds), *Sociological theories in progress*, Houghton Mifflin, Boston, MA.
- Doreian, P, Kapuscinski, R, Krackhardt, D and Szczypula, J 1996, A brief history of balance through time, *Journal of Mathematical Sociology*, 21.
- EBiz 2013, Top 15 Most Popular Websites | June 2013, viewed 1 June, 2013, <<http://www.ebizmba.com/articles/most-popular-websites>>.
- Ellison, N, Steinfield, C and Lampe, C 2006, Spatially Bounded Online Social Networks and Social Capital : The Role of Facebook, *International Communication Association*, 36.
- Ellison, N, Steinfield, C and Lampe, C 2007, The Benefits of Facebook “Friends:” Social Capital and College Students’ Use of Online Social Network Sites, *Journal of Computer-Mediated Communication*, 12, (4).
- Ellison, NB, Steinfield, C and Lampe, C 2011, Connection strategies: Social capital implications of Facebook-enabled communication practices, *New Media & Society*, 13, (6).
- Ellison, NB, Vitak, J, Gray, R and Lampe, C 2014, Cultivating Social Resources on Social Network Sites: Facebook Relationship Maintenance Behaviors and Their Role in Social Capital Processes, *Journal of Computer-Mediated Communication*, 19.
- El Emam, K 2000, A methodology for validating software product metrics, *National Research Council of Canada*, Ontario, Canada.
- Fenton, N 1994, Software measurement: a necessary scientific basis, *IEEE transactions on software engineering*, 20, (3).
- Fenton, N and Neil, M 2000, Software metrics: roadmap, Future of software engineering, Limerick, Ireland.
- Fukuyama, F 2001, Social capital , civil society and development, 22, (1).
- Garfield, S 2006, MVC vs. PAC, viewed <<http://www.garfieldtech.com/blog/mvc-vs-pac>>.
- Gargiulo, MMB 1999, The dark side of social capital’, *Social capital and liability*, Dordrecht: Kluwer.
- Gil de Zúñiga, H, Jung, N and Valenzuela, S 2012, Social Media Use for News and Individuals’ Social Capital, Civic Engagement and Political Participation, *Journal of Computer-Mediated Communication*, 17, (3).
- 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*, ACM Press, New York, New York, USA.
- Granovetter, M 1973, The strength of weak ties, *American Journal of Sociology*.

- Hagan, J, MacMillan, R and Wheaton, B 1996, New kid in town: social capital and the life course effects of family migration on children, *American Sociological Review*, 61.
- Hampton, KN 2003, Neighbouring in netville: how the internet supports community and social capital in a wired suburb, *City & Community*, 2, (4).
- Hanifan, LJ 1916, The Rural School Community Center, *American Academy of Political and Social Science*, 67.
- Heider, F 1946, Attitudes and Cognitive Orientation, *Journal of Psychology*, 21.
- Henderson, T, Hutton, L and Mcneilly, S 2012, Ethics and online social network research – developing best practices, *Proceedings of Proceedings of the 26th BCS Conference on Human Computer Interaction*.
- Hofer, M and Aubert, V 2013, Perceived bridging and bonding social capital on Twitter: Differentiating between followers and followees, *Computers in Human Behavior*, 29, (6), Elsevier Ltd.
- Ichiro, K, Kim, D, Coutts, A and Subrananian, S V 2004, Commentary: Reconciling the three accounts of social capital, *International Journal of Epidemiology*, 33.
- IEEE 1990, IEEE Standard 610.12-1990 Standard Glossary of Software Engineering Terminology.
- Junco, R 2012a, The relationship between frequency of Facebook use, participation in Facebook activities, and student engagement, *Computers & Education*, 58, (1).
- Junco, R 2012b, Too much face and not enough books: The relationship between multiple indices of Facebook use and academic performance, *Computers in Human Behavior*, 28, (1), Elsevier Ltd.
- Jung, Y, Gray, R, Lampe, C, Ellison, NB and Mi, AA 2013, Favors from Facebook Friends : Unpacking Dimensions of Social Capital, *CHUI3*, Paris, France.
- Kaplan, AM and Haenlein, M 2010, Users of the world, unite! The challenges and opportunities of Social Media, *Business Horizons*, 53, (1).
- Kitchenham, G, Pfleeger, L and Fenton, N 1995, Towards a framework for software measurement validation, *IEEE transactions on software engineering*, 21, (12).
- Krishna, A 2012, Active Social Capital: Tracing the roots of development and democracy, Columbia University Press, Columbia.
- Kumar, R and Stevenson, D 1997, CORRELATION COEFFICIENT: Critical values for Testing Significance, viewed <<http://empslocal.ex.ac.uk/people/staff/dbs202/cat/stats/corr.html>>.
- Lewis, K, Kaufman, J, Gonzalez, M, Wimmer, A and Christakis, N 2008, Tastes , ties , and time : A new social network dataset using Facebook . com, *Social Networks*, 30.

- Liccardi, I, Ounnas, A, Pau, R, Massey, E, Kinnunen, P, Lewthwaite, S, Midy, M-A and Sakar, C 2007, The role of social networks in students' learning experiences, *Working group reports on ITiCSE on Innovation and technology in computer science education ITiCSEWGR 07*, 39, (4), ACM.
- Lin, N 2008, A network theory of social capital, in D Castiglione, JW van Derth & G Wolleb (eds), *The handbook of social capital*, Oxford University Press, London.
- Lin, N 1999, Building a Network Theory of Social Capital, 22, (1).
- Lin, N 2001, *Social Capital*, Cambridge University Press, Cambridge, UK.
- Loury, G 1977, A dynamic theory of racial income differences in Women, Minorities, and Employment Discrimination, *AM La Mond*, MA: Heath, Lexington.
- MacLanahan, S and Sandefur, G 1994, Growing up with a single parent what helps what hurts, Harvard University Press, Cambridge MA.
- Martino, F and Spoto, A 2006, Social Network Analysis : A brief theoretical review and further perspectives in the study of Information Technology, *PsychNology Journal*, 4, (1).
- Marx, K 1867, *Capital: Volume I, Der Produktionsprozess Des Kapitals*, Verlag von Otto Meisner, Germany.
- McCalla, G 2000, The fragmentation of culture, learning, teaching and technology, *International journal of Artificial intelligence*, 11, (12).
- Meneely, A 2012, Validating Software Metrics : A Spectrum of Philosophies, *ACM Transactions on Software Engineering and Methodology (TOSEM)*, 21, (4).
- Mislove, A, Marcon, M, Gummadi, KP, Druschel, P and Bhattacharjee, B 2007, Measurement and analysis of online social networks, *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement - IMC '07*, ACM Press, New York, New York, USA.
- Moody, J 2001, Race, school integration, and friendship segregation in America, *American journal of sociology*, 107.
- Moody, J and Paxton, P 2009, Building bridges: Linking Social Capital and Social Networks to improve theory and research, *American Behavioral Scientist*, 52, (11).
- Moreno, J 1946, Sociogram and sociomatrix, *Sociometry*, 9.
- Moreno, J 1934, *Who shall survive?*, Beacon Press, New York, USA.
- Moreno, M, Fost, N and Christakis, D 2008, Research ethics in the MySpace era, *Pediatrics*, 121.

- Nahapiet, J and Ghoshal, S 1998, Social capital, intellectual capital, and the organizational advantage, *Academy of management review*, 23, (2).
- Narayan, D and Pritchett, L 1997, Cents and sociability: Household income and social capital in rural Tanzania, World Bank, Washington DC, USA.
- Nodexl.com 2014, NodeXL.
- Norris, P 2002, The bridging and bonding role of online communities, *Harvard International Journal of Press/Politics*, 7, (3).
- OECD 2001, The well-being of nations: the role of human and social capital, *Observer*, 226/227.
- Opsahl, T and Panzarasa, P 2009, Clustering in Weighted Networks, *Social Networks*, 31, (2).
- Page, L, Brin, S, Motwani, R and Winograd, T 1999, The PageRank citation ranking: bringing order to the web.
- Pfeil, U, Arjan, R and Zaphiris, P 2009, Age differences in online social networking – A study of user profiles and the social capital divide among teenagers and older users in MySpace, *Computers in Human Behavior*, 25, (3), Elsevier Ltd.
- Pfeil, U, Arjan, R and Zaphiris, P 2009, Age differences in online social networking – A study of user profiles and the social capital divide among teenagers and older users in MySpace, *Computers in Human Behavior*, 25, (3), Elsevier Ltd, pp. 643–654.
- Portes, A 1998, Social Capital: Its Origins and Applications in Modern Sociology, *Annual Review of Sociology*, 24, (1).
- Portes, A 2000, The Two Meanings of Social Capital, 15, (1).
- Prusak, L and Cohen, D 2001, In good company: How social capital makes organisations work, Harvard business school press, Boston, MA.
- Putnam 1993, Making democracies work, Princeton University Press, Princeton.
- Putnam, R 2000, Bowling Alone: The Collapse and Revival of American Community, Simon & Schuster, New York, USA.
- Putnam, R 1995, Tuning in, tuning out; the strange disappearance of social capital in America, *Political sciences and politics*.
- Rafaeli, S, Ravid, G and Soroka, V 2004, De-lurking in virtual communities : a social communication network approach to measuring the effects of social and cultural capital, *Proceesings of the 37th annual Conference IEEE Hawaii*.
- Rao, P 2007, Statistical research methods in life sciences, Duxbury Press, Pacific grove CA.

- Reenskaug, T 2003, The Model-View-Controller (MVC) Its Past and Present, viewed <http://heim.ifi.uio.no/~trygver/2003/javazone-jaoo/MVC_pattern.pdf>.
- Resnick, P 2001, Beyond Bowling Together: SocioTechnical Capital Activity Social Capital Activity, *HCI in the new millenium*, (March).
- Resnick, P 2004, Impersonal Sociotechnical Capital , ICTs , and Collective Action Among Strangers, *Transforming*.
- Roche, J 1994, Software metrics and measurement principles, *ACM SIGSOFT Software Engineering*, 19, (1).
- Scheufele, DA and Shah, D V 2000, Personality Strength and Social Capital, 27, (2).
- Schiff, M 1992, Social capital, labor mobility, and welfare, *Rational Society*, 4.
- Schofield, J 1979, The impact of positively structured contact on intergroup behaviour: does it last under adverse conditions, *Social Psychology Quarterly*, 42.
- Smith, DK 2003, Shortest Paths, *Princeton*, viewed <<http://www.cs.princeton.edu/~rs/AlgsDS07/15ShortestPaths.pdf>>.
- Smith, MS 2008, Social Capital in Online Communities, *Simulation, Computing Methodologies*.
- Solberg, L 2010, Data mining on Facebook: A free space for researchers or an IRB nightmare, *Univerisity of Illionois Journal of Law, Tehcnology & Policy*.
- Soper, D 2014, p-Value Calculator for Correlation Coefficients, viewed <<http://www.danielsoper.com/statcalc3/calc.aspx?id=44>>.
- Stangroom, J 2014, Social Science Statistics, viewed <<http://www.socscistatistics.com/pvalues/pearsondistribution.aspx>>.
- Steinfeld, C, Dimicco, JM, Ellison, NB and Lampe, C 2009, Bowling Online : Social Networking and Social Capital within the Organization, *C & T 09*, ACM, Pennsylvania.
- Steinfeld, C, Ellison, NB and Lampe, C 2008, Social capital, self-esteem, and use of online social network sites: A longitudinal analysis, *Journal of Applied Developmental Psychology*, 29, (6), Elsevier B.V.
- Sun, Y and Shang, R-A 2014, The interplay between users' intraorganizational social media use and social capital, *Computers in Human Behavior*, 37, Elsevier Ltd.
- Tamaschke, L 2003, The Role of Social Capital in Regional Technological Innovation : Seeing both the wood and the trees, *Internatinal conference on communities and technolgies*, Amsterdam.
- Tsai, W and Ghoshal, S 1998, Social capital and value creation, *Academy of Management Journal*, 41, (4).

- Twittercounter.com 2014, Twitter Top 100 Most Followers, *Twittercounter.com*, viewed <<http://twittercounter.com/pages/100>>.
- Valente, TW 2010, *Social Networks and Health*, Oxford University Press, New York.
- Valenzuela, S 2008, Lessons from Facebook: The Effect of Social Network Sites on College Students' Social Capital, *Symposium on online journalism*, Austin Texas.
- Valenzuela, S, Park, N and Kee, KF 2009, Is There Social Capital in a Social Network Site?: Facebook Use and College Students' Life Satisfaction, Trust, and Participation, *Journal of Computer-Mediated Communication*, 14, (4).
- Wang, Y-C, Burke, M and Kraut, R 2013, Gender, topic and audience response: an analysis of user-generated content on Facebook, *CHI 13: Proceedings of the SIGCHI conference on human factors in computing systems*.
- Wasko, MM and Faraj, S 2005, Why should I share? Examining social capital and knowledge contribution in electronic networks of practice, *MIT Quarterly*, 29, (1).
- Wasserman, S and Faust, K 1994, *Social network analysis: Methods and applications*, Cambridge University Press, Cambridge.
- Wellman, B 1997, An electronic group is virtually a social network, in S Kiesler (ed.), *Culture of the Internet*, Lawrence Erlbaum, New Jersey.
- Wellman, B and Frank, K 2001, Network capital in a multi-level world: Getting support from personal communities, in N Line, K Cook & R Burt (eds), *Social capital: Theory and research*, Hawthorne, New York, USA.
- Wellman, B, Haase, AQ, Witte, J and Hampton, K 2001, Does the Internet Increase , Decrease , or Supplement Social Capital ?, 45, (3).
- Willging, PA 2005, Using Social Network Analysis Techniques to Examine Online Interactions, 2, (9).
- Williams, D 2006, On and off the 'net: Scales for social capital in an online era, *Journal of Computer-Mediated Communication*, 11, (2).
- Woolcock, M 1998, Social capital and economic development : Toward a theoretical synthesis and policy framework, *Theory and Society*, 27, (2).
- Woolcock, M 2001, The place of social capital in understanding social and economic outcomes, *Isima: Canadian Journal of Policy Research*, 2, (1).
- WorldBank 1999, Social Capital Research Group, viewed <<http://go.worldbank.org/K4LUMW43B0>>.
- Yoder, C, Hill, C and Stutzman, F 2011, Identifying Social Capital in the Facebook Interface.

Zhao, S 2006, Do Internet Users Have More Social Ties? A Call for Differentiated Analyses of Internet Use, *Journal of Computer-Mediated Communication*, 11, (3).

Appendix A Internet Social Capital Scales

Bonding Subscale

1. There are several people online/offline I trust to help solve my problems.*
2. There is someone online/offline I can turn to for advice about making very important decisions.*
3. There is no one online/offline that I feel comfortable talking to about intimate personal problems. (reversed)*
4. When I feel lonely, there are several people online/offline I can talk to.
5. If I needed an emergency loan of \$500, I know someone online/offline I can turn to.*
6. The people I interact with online/offline would put their reputation on the line for me.
7. The people I interact with online/offline would be good job references for me.
8. The people I interact with online/offline would share their last dollar with me.
9. I do not know people online/offline well enough to get them to do anything important. (reversed)
10. The people I interact with online/offline would help me fight an injustice.

Bridging Subscale

1. Interacting with people online/offline makes me interested in things that happen outside of my town.
2. Interacting with people online/offline makes me want to try new things.
3. Interacting with people online/offline makes me interested in what people unlike me are thinking.
4. Talking with people online/offline makes me curious about other places in the world.
5. Interacting with people online/offline makes me feel like part of a larger community.
6. Interacting with people online/offline makes me feel connected to the bigger picture.
7. *Interacting with people online/offline reminds me that everyone in the world is connected.*
8. I am willing to spend time to support general online/offline community activities.
9. Interacting with people online/offline gives me new people to talk to.
10. *Online/Offline, I come in contact with new people all the time.*

(Williams 2006)

Appendix B Preliminary Analysis Interview & Survey Questions

The results from the interviews (anonymised) –

http://sociallearningspace.abertay.ac.uk/scism/Interview_Questions_Summary_2012_anonymised.doc

Summary of Interview question answers

General

Academic year you joined the university: first, second etc?

Second

Second

Second

First - mature student

First

First

First

Second

Did you go to college, if so where and what did you study?

Angus College

Angus College

Angus College

Exeter University

NA

Falkirk

School

Dundee college

When you first came to the University where there any particular challenges or problems that you found such as finding information, people or places?

Ok

OK, info was fine, making friends ok

Would have liked info on coursework load, teaching style

Getting around would be nice

Finding places was difficult

Ok – things were pretty easy

People and places info

Ok

Social Connections

Roughly how many offline friends/acquaintances do you have? Broadly speaking this is people that you know.

30

50

60

80

30

25

550

30

Roughly how many online friends/acquaintances do you have? Broadly speaking this is people that you communicate with online.

300
500
190
40 (-)
70
100
350
355

Is you had a task such as finding out more about say a book, film or music, what strategy would you most often adopt? Your answer may refer to anything such as offline, google, social media or any combination.

Amazon, google
Google, facebook
Google,
Googe and review
Google maybe review app
Google may SN
Amazon- ? sm? Word of mouth
Google - magazines

Is you had a task such as finding out more about an academic topic that you were studying, what strategy would you most often adopt? Your answer may refer to anything such as offline, google, social media or any combination.

Google, w3schools
Google, facebook
Blackboard, google
Google check sources
Library, ebooks, google
Google maybe SN grp
Virtual library, goolge
googel

What if any social media do you use?

Facebook, Twitter used to use myspace and bebo
Facebook, google plus, twitter, messaging
FB
Fb, twitter, linkin
FB occasional twitter
FB
Fb, twitter,
FB

Is making social communication a goal of using social media, or do you have something else in mind such as entertainment?

Drivers – comms and entertainment
Communication channel
Comms.
Use FB because everyone uses it
Other user effect ? number and opp?
Communications
Comms

Communication

In order of rank how do you communicate most often do you use facebook, sms, email or other?

SMS, FB only uni email
FB, sms occasionally, Email mainly for university
SMS, FB only uni email
Email, sms
Text, fb messages
SMS, FB only uni email
Sms, email, fb
Facebook, facebook chat, sms

In order of rank what types of things do you talk about when using social media?

What you are doing, where you are, 1,1,1,3, 1,0,0,1
Sharing information, 2,0,3,1, 3, 1,0,2
Entertainment: films, books, music, 3 ,1,0,2,3,0,1,4
Academic or University life 1=,0,2,4,2, 2,2,3
Work 4,0,4,5, 4,0,0,5
other?

What types of things influences whether or not you decide to post to social media?

Relevance only
Privacy concerns against things like tagging
Privacy against comms as a driver
Access to group ask
Academic enquiry of interest
Boredom, distraction
Friends, interests
Answer a question

Facebook

How long have you used FB it and often do you use it?

2008, several times a day
2009 over ten times a day
2009, several times a day
2008, several times a day
2008, 3-4 times a day
2009, 2-3 times a day
2008, several time a day
2009, 1-2 a day

Do you use it to either view or post content most often or about the same?

view
view
view
view
view
view
view
view

Do you communicate using social media with people that you have not met offline? If so how often?

yes
yes
not really
yes on twiter but not FB
not really
No
Occasionally
Not really

Do you use FB groups, WDD or other?

Yes - Wdd, ebay group
Yes - Wdd
Yes – Wdd, music
Yes – Wdd, hobby
Yes – wdd, oop, ce0732a and friends
Yes – WDD and for group module
Yes – WDD, family, church
Yes - wdd

How often do you view/post FB groups in general?

Daily – is influenced by the notification prompts
Occasional – influenced by social and events
Occasional
View several times a day, only occasional posting
Several times a day
Several times a day – when you see the prompt
Most recent most of 30 mins?
Daily occasionaly

Is whether or not you post to an FB group affected by topic?

Topic
Topic – question wants answer (roles in groups are ok)
Yes – happy to help but wary of giving code away
Yes
Topic
Topic
Topic
Yes mainly

If you saw a post from a friend are you more or less likely to comment?

No difference
Not really
No
More likely to a friend
Yes to a friend
Yes
Not really
More likely

In regard to whether or not to comment on a post are you influenced by how many comments a post has?

No
No – topic
No – content
Perhaps – more likely answer a question no one else has

Maybe
No difference
Not really
Sometimes – more likely to post to where there are only a few comments

Is there anything that might influence you to comment on a post from a person that you did not know? If topic and then other than topic.

Topic
Phrasing
Rely???
Not many comments
Topic
Mainly topic
Mainly topic
Topic

If you knew that a person did not have a great deal of comments and wanted more would you be persuaded by a prompt from the software, of some kind, to comment?

Maybe, 'priority posts'
Yes, 'newcomer'
Yes – no to newcomer, yes to priority posts
Yes unanswered posts
Yes to buttons for priority posts
Yes – newcomer, roles for members
Maybe – mainly interest matching
Not sure

The results from the surveys (anonymised) –

http://sociallearningspace.abertay.ac.uk/scism/Facebook_Survey_2012_anonymised.xlsx

1	Timestamp	Do you have a facebook account	Approximately how many friends do you have?	Are you currently a member of the VDD facebook group?	To what extent do the following terms describe your opinion of VDD facebook group? [The VDD facebook group has helped me with my general knowledge]	To what extent do the following terms describe your opinion of VDD facebook group? [The VDD facebook group helped me find specific answers to]	To what extent do the following terms describe your opinion of VDD facebook group? [The VDD facebook groups helped me make friends]	To what extent do the following terms describe your opinion of VDD facebook group? [The VDD facebook group helps build a sense of community with others on]	On average how often do you check the VDD facebook group?	Overall how would you describe the involvement of academic staff on the VDD facebook group?
14	4/28/2012 12:19:47	Yes	Yes	80 No						
15	4/28/2012 12:32:38	Yes	Yes	13 Yes	Strongly Agree	Strongly Agree	Neutral	Agree	More than 10 times	Positive
16	4/28/2012 13:46:11	Yes	Yes	50 Yes	Agree	Neutral	Neutral	Strongly Agree	2-5 times a day	Positive
17	4/28/2012 14:52:04	Yes	Yes	100 No						
18	4/28/2012 16:53:57	Yes	Yes	200 No						
19	4/29/2012 0:24:24	Yes	Yes	150 Yes	Strongly Disagree	Neutral	Neutral	Agree	Once a day	Neutral
20	4/29/2012 9:20:35	Yes	Yes	50 No						
21	4/29/2012 11:16:45	Yes	Yes	500 No						
22	4/29/2012 14:01:36	Yes	Yes	15 Yes	Agree	Neutral	Neutral	Strongly Agree	2-5 times a day	Very Positive
23	4/29/2012 15:25:14	Yes	Yes	250 No						
24	4/29/2012 20:47:22	Yes	Yes	110 Yes	Neutral	Agree	Disagree	Agree	Less than once a	Positive
25	4/30/2012 10:58:00	Yes	Yes	50 No						
26	4/30/2012 14:21:39	Yes	70-80	Yes	Agree	Agree	Agree	Agree	Once a week	Positive
27	5/1/2012 16:51:05	Yes	Yes	30 Yes	Disagree	Agree	Agree	Strongly Agree	2-5 times a day	Very Positive
28	5/2/2012 11:00:29	Yes	Yes	50 Yes	Agree	Agree	Agree	Strongly Agree	Once a day	Neutral
29	5/2/2012 15:31:14	Yes	Yes	100 No						
30	5/7/2012 13:59:55	Yes	Yes	30 Yes	Agree	Strongly Agree	Strongly Disagree	Strongly Agree	Once a day	Very Positive
31	5/9/2012 16:15:29	Yes	Yes	10 Yes	Neutral	Neutral	Disagree	Agree	Once a day	Very Positive
32	5/9/2012 16:28:20	Yes	Yes	15 Yes	Strongly Agree	Agree	Agree	Strongly Agree	Once a week	Very Positive
33	5/9/2012 16:50:18	Yes	Yes	20 Yes	Agree	Neutral	Neutral	Agree	2-5 times a day	Neutral
34	5/9/2012 19:07:41	Yes	Yes	80 Yes	Agree	Agree	Disagree	Agree	Once a day	Positive
35	5/9/2012 19:28:05	Yes	Yes	300 No						
36	5/9/2012 19:37:51	Yes	Yes	80 Yes	Strongly Disagree	Strongly Disagree	Strongly Disagree	Neutral	Once a day	Positive
37	5/9/2012 19:37:53	Yes	Yes	25 Yes	Strongly Agree	Neutral	Strongly Agree	Strongly Agree	Once a day	Positive
38	5/9/2012 23:46:03	Yes	Yes	20 Yes	Neutral	Neutral	Strongly Disagree	Disagree	Once a day	Positive
39	5/10/2012 10:54:25	Yes	Roughly 25	Yes	Agree	Disagree	Agree	Neutral	2-5 times a day	Neutral
40	5/10/2012 10:55:58	Yes	Yes	12 Yes	Disagree	Neutral	Disagree	Strongly Disagree	Once a day	Neutral
41	5/10/2012 13:01:41	Yes	Yes	10 Yes	Agree	Strongly Agree	Neutral	Strongly Agree	Once a day	Positive
42	5/10/2012 23:20:30	Yes	Yes	100 Yes	Neutral	Agree	Disagree	Neutral	Once a week	Positive
43										
44										

Timestamp	On average how many times a week do you check WebCT?	Please leave a comment on your experience of facebook for University related purposes and/or make suggestions for next year.	Please indicate to what extent you agree or disagree with the following descriptive terms about your impression of the VDD facebook	Please indicate to what extent you agree or disagree with the following descriptive terms about your impression of the VDD facebook	Please indicate to what extent you agree or disagree with the following descriptive terms about your impression of the VDD facebook	Please indicate to what extent you agree or disagree with the following descriptive terms about your impression of the VDD facebook	On average how often do you post or comment to the VDD facebook group?
4/28/2012 12:19:47	Once a day	I have been invited					
4/28/2012 12:32:38	Once a day	Almost every post	Strongly Disagree	Strongly Agree	Strongly Agree	Agree	Once a day
4/28/2012 13:46:11	Once a day	Facebook groups	Agree	Disagree	Strongly Agree	Agree	Never
4/28/2012 14:52:04	Less than once a						
4/28/2012 16:53:57	Once a day						
4/29/2012 0:24:24	Once a week	The 'VDD' is good	Neutral	Neutral	Disagree	Neutral	Less than once a
4/29/2012 9:20:35	Once a week						
4/29/2012 11:16:45	Less than once a						
4/29/2012 14:01:36	Once a day	I think the VDD	Strongly Agree	Disagree	Agree	Agree	Less than once a
4/29/2012 15:25:14	Once a day	Facebook is pretty					
4/29/2012 20:47:22	2-5 times a day		Neutral	Neutral	Strongly Agree	Neutral	Less than once a
4/30/2012 10:58:00	Once a week	Very useful.					
4/30/2012 14:21:39	2-5 times a day		Neutral	Neutral	Neutral	Agree	Less than once a
5/1/2012 16:51:05	Less than once a	The idea of	Strongly Agree	Disagree	Agree	Agree	Less than once a
5/2/2012 11:00:29	Once a day	As far as am was	Agree	Strongly Disagree	Agree	Strongly Agree	Once a week
5/2/2012 15:31:14	More than 10 times						
5/7/2012 13:59:55	Less than once a	Facebook is an	Neutral	Strongly Disagree	Strongly Agree	Strongly Agree	Once a week
5/9/2012 16:15:29	Once a day		Agree	Strongly Disagree	Strongly Agree	Neutral	Never
5/9/2012 16:28:20	Once a week	The 'VDD' group is	Agree	Neutral	Strongly Agree	Agree	Less than once a
5/9/2012 16:50:18	Less than once a		Agree	Agree	Agree	Strongly Agree	Once a week
5/9/2012 19:07:41	Once a day		Neutral	Agree	Agree	Neutral	Less than once a
5/9/2012 19:28:05	Once a week	I would rather to					
5/9/2012 19:37:51	Less than once a	The majority of	Neutral	Neutral	Disagree	Strongly Disagree	Less than once a
5/9/2012 19:37:53	Once a day		Strongly Agree	Strongly Disagree	Strongly Agree	Agree	Less than once a
5/9/2012 23:46:03	Once a day	It needs to be	Agree	Neutral	Agree	Strongly Agree	Less than once a
5/10/2012 10:54:25	Once a day	In regards to the	Disagree	Agree	Neutral	Strongly Agree	Once a week
5/10/2012 10:55:58	Less than once a	I joined 'VDD' as it	Neutral	Strongly Agree	Disagree	Neutral	2-5 times a day
5/10/2012 13:01:41	Once a day	I have found	Agree	Agree	Strongly Agree	Disagree	Less than once a
5/10/2012 23:20:30	2-5 times a day	I think the	Agree	Disagree	Agree	Neutral	Never

	A	B	C	D	E	F	G
		On average how many times a week do you check WebCT?	Approximately how many acquaintances do you have?	Approximately how many, if any, friends and acquaintances do you think that you have added as a result of	Approximately how many facebook friends do you have?	Approximately how many, if any, facebook friends do you think that you have added as a result of studying at Abertag?	
1	Timestamp						
14	4/28/2012 12:19:47	Once a day	80	25	298	20	
15	4/28/2012 12:32:38	Once a day	20/30	3	245	8	
16	4/28/2012 13:46:11	Once a day	50	20	125	10	
17	4/28/2012 14:52:04	Less than once a	200	40	300	20	
18	4/28/2012 16:53:57	Once a day	20	10	250	20	
19	4/29/2012 0:24:24	Once a week	60	5	220	5	
20	4/29/2012 9:20:35	Once a week	100	100	150	100	
21	4/29/2012 11:16:45	Less than once a	150	200	640	50	
22	4/29/2012 14:01:36	Once a day	40	20	43	5	
23	4/29/2012 15:25:14	Once a day	20	15	327	20	
24	4/29/2012 20:47:22	2-5 times a day	150	20	200	20	
25	4/30/2012 10:58:00	Once a week	100	10	160	8	
26	4/30/2012 14:21:39	2-5 times a day	80	7/6/2012	67	7/6/2012	
27	5/1/2012 16:51:05	Less than once a	5	3	35	3	
28	5/2/2012 11:00:29	Once a day	30	45	200	40	
29	5/2/2012 15:31:14	More than 10 times	85	10	464	10	
30	5/7/2012 13:59:55	Less than once a	100	30	580	100	
31	5/9/2012 16:15:29	Once a day	20	5	120	18	
32	5/9/2012 16:28:20	Once a week	25	40	423	40	
33	5/9/2012 16:50:18	Less than once a	60	30	231	50	
34	5/9/2012 19:07:41	Once a day	30	20	173	20	
35	5/9/2012 19:28:05	Once a week	20	4	100	4	
36	5/9/2012 19:37:51	Less than once a	650	60	350	40	
37	5/9/2012 19:37:53	Once a day	50	10	600	30	
38	5/9/2012 23:46:03	Once a day	30	40	70	40	
39	5/10/2012 10:54:25	Once a day	50	7	536	7	
40	5/10/2012 10:55:58	Less than once a	300	14	350	14	
41	5/10/2012 13:01:41	Once a day	30	10	350	10	
42	5/10/2012 23:20:30	2-5 times a day	200	100	280	100	

Appendix C SCiSM Software Code

```
<?php

/*
    SCiSM php script - 2014
    All rights reserved Abertay University
*/

ini_set("display_errors", 1);
error_reporting(E_ALL);

/////////////////////////////////CONTROLLER/////////////////////////////////

class Controller{
    // XML data array
    public $start = array();

    // file error handling

    /* READ XML INTO DATA ARRAY
    make 'data' an associative array for 'record' elements, with the keys
    'left', 'right',
    'messageId' and 'weekNumber'
    individual values in the array will be accessed using
    data[row_number]['some_key'] */
    public function fetchData() {

        $dataSource = "data/test.xml";

        $this -> start = json_decode(json_encode((array)
simplexml_load_file($dataSource)),1);

        $this -> start = $this ->start['record'];

        // Exit if file load returns false
        if ($this -> start==0)
        {
            exit( "<br> FILE LOAD FAILURE - ABORTED ");
        }

        // unlink($dataSource);
    }
}

/////////////////////////////////MODEL/////////////////////////////////

class Model{

    public $results = array();

    // filter the XML data so that it contains only the specified column,
    // then return how many times the subject appears in the resulting
data
    private function processArray($subject, $column, $start) {

        $filtered = array();
        $counted = array();
```

```

        foreach($start as $row) {
            $filtered[count($filtered)] = $row[$column];
        }

        $counted = array_count_values($filtered);
        return $counted[$subject];
    }

    // calculates out-degree of subject in left column or return 0
    private function calculateOutDegree($subject, $start, $col) {

        if ($col=="right"){
            return 0;
        }
        // return calculation to be written to array
        return $this->processArray($subject, "left", $start);
    }

    // calculates in-degree of subject
    private function calculateInDegree($subject, $start) {
        // return calculation to be written to array
        return $this-> processArray($subject, "right", $start);
    }

    // calculates degree of subject
    private function calculateDegree($subject, $start, $col) {
        // return calculation to be written to array
        $answer = $this->calculateInDegree($subject, $start)+$this-
>calculateOutDegree($subject, $start, $col);
        return $answer;
    }

    // finds subject in array
    private function findTarget($filtered, $countedseen, $j, $subject){

        foreach ($filtered as $rel){

            if ( ($rel["left"]==$subject)&&
($rel["right"]==$countedseen[$j]) )
            {

                return 1;
            }

        }

        }//end foreach

        return 0;

    }//end function

    // calculates clustering coefficient
    private function calculateCC($subject, $start) {
        // filter the data so that it contains only the specified
column,
        // then return how many times the subject appears in the
resulting data
        $filtered = array();
        $counted = array();
        $countedseen = array();
        $countednotseen = array();

```

```

$hit=0;
$miss=0;

//get rid of duplicates in the array (may need debugging0
doesnt seem to effect the result
//$filtered = array_map("unserialize",
array_unique(array_map("serialize", $start)));
//$filtered=array_values(array_unique($start));
$filtered=$start;

//get the links in both left and right columns
foreach ($filtered as $know)
{
    if( ($know["left"]== $subject) &&
($know["right"]!= $subject)){
        $counted [count($counted)] = $know["right"];
    }

    if(( $know["right"]== $subject) &&
($know["left"]!= $subject)){
        $counted [count($counted)] = $know["left"];
    }
}
//get rid of duplicates in the array and tidy the array index
$counted=array_values(array_unique($counted));

$counted=count($counted);

//split the counted array into two arrays one containing seen
the other the rest
foreach ($counted as $rows)
{
    if(strpos($rows, 'Seen') === 0)
    {
        $countedseen[count($countedseen)] =$rows;
    }
    else
    {
        $countednotseen [count($countednotseen)] =$rows;
    }
}

for ($j=0;$j<count($countedseen);$j++)
{
    $subjectfinder=0;
    $monitor=$this->findTarget($filtered, $countedseen,
$j, $subject);
    if ($monitor ==1)

```

```

        {
            $hit++;
        }
        else if ($monitor ==0)
        {
            $miss++;
        }
    } //end for

    if ($hit!=0)
    {
        $answer = $hit/($hit+$miss);

        return $answer;
    }
    else
    {
        return 0;
    }
}

// counts the number of times a message is seen for SCiSM calculation
private function countSeen($start){
    $seenCount=0;
    $column="left";
    foreach($start as $row) {

        if(strpos($row[$column],'Seen') === 0)
        {
            $seenCount++;
        }
    }

    return $seenCount;
}

// calculates bonding SCiSM using out-degree and clustering
coefficient
private function calculateBondingScism($subject, $start, $col)
{
    return $this->calculateOutDegree($subject, $start, $col)*$this-
>calculateCC($subject, $start);
}

// calculates bridging SCiSM using in-degree and count seen (number
of times the messages is seen)
private function calculateBridgingScism($subject, $start)
{
    return round($this->calculateInDegree($subject, $start)/$this-
>countSeen($start)*10,4);
}

// calculates total SCiSM score

```



```

        right column i.e. commentors or viewers */
class View{

    public function displayResults($results, $type) {

        $html = "<h2>Contents of Results Array".$type."</h2>";

        // construct HTML table
        $html .= "<table>";
        $html .= "<thead>";
        $html .= "<tr>";

        $html .=
        "<th>weekNumber</th><th>subject</th><th>degree</th><th>inDegree</th><th>out
        Degree</th><th>Cluster
        Coef.</th><th>bondingScism</th><th>bridgingScism</th><th>totalScism</th>";

        $html .= "</tr>";
        $html .= "</thead>";
        $html .= "<tbody>";

        for($i=0;$i<count($results);$i++) {
            $html .= "<tr>";

            $html .= "<td>".$results[$i]["weekNumber"]."</td>";
            $html .= "<td>".$results[$i]["subject"]."</td>";
            $html .= "<td>".$results[$i]["degree"]."</td>";
            $html .= "<td>".$results[$i]["inDegree"]."</td>";
            $html .= "<td>".$results[$i]["outDegree"]."</td>";
            $html .= "<td>".$results[$i]["cc"]."</td>";
            $html .= "<td>".$results[$i]["bondingScism"]."</td>";
            $html .= "<td>".$results[$i]["bridgingScism"]."</td>";
            $html .= "<td>".$results[$i]["totalScism"]."</td>";

            $html .= "</tr>";
        }

        $html .= "</tbody>";
        $html .= "</table> <br> <br>";

        echo $html;
    }

}

} //end view class

//////////////////////////////////INSTANTIATIONS//////////////////////////////////

/* The objects are instantiated and calls made
   this works by chaining the data arrays between objects */

// load xml in to 'start' data array
$control = new Controller();
$control -> fetchData();

// calculate 'start' data into 'results' array for posters i.e.
'left' in xml column
$model = new Model();
$model-> calculateResults($control->start, "left");

```

```
// displays 'results' array and title
$view = new View();
$view-> displayResults($model -> results, " subjects with out-
degree");

// calculate 'start' data into 'results' array for commenters and
viewers i.e. 'right' in xml column
$model-> calculateResults($control->start, "right");

// displays 'results' array and title
$view-> displayResults($model -> results, " all subjects in-degree
only ");

/*
Instructions for updating software

1. Controller - Alter .xml file for input
2. View - add column to the display
3. Model - add column to the calculate results method
4. Model - add method to calculate new results

*/
?>
```

Appendix D Data preparation for Data Set 1

The data from the discussion forum was written in to an XML file which was then parsed by the SCiSM software. The steps taken in preparing the XML data are summarised as follows:

People in the message

Left right

Poster> commenter

People in the message

This is commenter back to poster IF poster is below commenter in the thread

Eg week 3

Left right

Commentor>Poster

People in the message

This is commenter back to another commenter IF refers to previous comment e.g week 4

Left right

Commentor>commentor

Poster and commenter (for every message)

Left right

Person > messageid

Non poster or commenter aka seen

Left right

Messageid> person

Poster and commenters

See rule 3 are also seen

These arrangements are illustrated over the page.

Illustration of XML preparation

	A	B	C
1			Visual Prope
2	Vertex 1	Vertex 2	Color
3	James Dunt	Lynsay She	Blue
4	James Dunt	James Dunt	Blue
5	James Dunt	James Dunt	Blue
6	Lynsay She	James Dunt	Blue
7	James Dunt	SeenJP	Blue
8	Lynsay She	SeenJP	Blue
9	James Dunt	SeenJP	Blue
10	James Dunt	SeenJP	Blue
11	SeenJP	James Dunt	Blue
12	SeenJP	Adam Elbuzedi	
13	SeenJP	Adrian Robb	
14	SeenJP	Aimi Forgan	
15	SeenJP	Alex Pantelji	
16	SeenJP	Alkyum Ash	
17	SeenJP	Amy McLeod	
18	SeenJP	Andrei Jules Paul	
19	SeenJP	Callum Hutchison	
20	SeenJP	Chris Murdoch	
21	SeenJP	Christian Pauli	
22	SeenJP	Craig Mitchell	
23	SeenJP	Craig Stokes	

Poster > Comment

Poster > Seen

Seen > Poster

Comment > Seen

Seen > Comment

Comment > Poster

	A	B	C	D
1				Visual Properties
2	Vertex 1	Vertex 2	Color	Width
81	SeenJP	Stephen Webster		
82	SeenJP	William Cameron		
83	Neal Mulle	Darryl Hutt	Blue	
84	Neal Mulle	Neal Mulle	Blue	
85	Neal Mulle	Mohamme	Blue	
86	Darryl Hutt	Neal Mulle	Blue	
87	Neal Mulle	SeenNM	Blue	
88	Darryl Hutt	SeenNM	Blue	
89	Neal Mulle	SeenNM	Blue	
90	Mohamme	SeenNM	Blue	
91	SeenNM	Neal Mulle	Blue	
92	SeenNM	Adam Elbuzedi		
93	SeenNM	Adrian Robb		
94	SeenNM	Aimi Forgan		
95	SeenNM	Alex Pantelji		
96	SeenNM	Alkyum Ash		
97	SeenNM	Amy McLeod		
98	SeenNM	Andrei Jules Paul		
99	SeenNM	Blaine McGowan		
100	SeenNM	Callum Hutchison		
101	SeenNM	Chris Murdoch		

Like is Poster > Like AND Like > Seen

Rationale

Poster > Comment – information sent and received as evidenced by reply

Poster . Seen – connection to the seen object on the graph for a weak tie

Seen >Poster - connection to the seen object on the graph as reciprocal – represent conversation

Comment > Seen - connection to the seen object on the graph for weak tie

Comment > Poster – where poster is below comment to indicate sent and received

Appendix E ISCS for Facebook Groups

The Facebook adapted ISCS questions were verified by the interviewees in the preliminary analysis. The survey can be viewed at

<https://docs.google.com/forms/d/1SOJcxTs9mRFYOTrsOwF8niaoIcPxgirmuPOJf7xfeCQ/viewform> and below

ISCS Social Capital Survey

This is a survey intended to examine the social value that users obtain from Facebook.

The answers that you give will be treated confidentially. Your name will never be mentioned in the reporting of the results.

Participation is entirely voluntary and you don't have to answer all of the questions.

It will take approximately five minutes to complete the questions.

Facebook Social Value

Please answer these questions with particular reference to how you have felt about the software over the last week.

1. There are several people on Facebook that I trust to help solve problems

1 2 3 4 5

Strongly Disagree Strongly Agree

2. There is someone on Facebook I can turn to for advice about making very important decisions

1 2 3 4 5

Strongly Disagree Strongly Agree

3. There is someone on Facebook that I feel comfortable talking to about personal problems

1 2 3 4 5

Strongly Disagree Strongly Agree

4. If I felt lonely, there are several people on Facebook that I could talk to

1 2 3 4 5

Strongly Disagree Strongly Agree

5. If I needed to borrow something of value, I know someone on Facebook I can turn to.

1 2 3 4 5

Strongly Disagree Strongly Agree

Continued over page

6. The people I interact with on Facebook would stick up for me.

1 2 3 4 5

Strongly Disagree Strongly Agree

7. The people I interact with on Facebook would speak well of me if asked to.

1 2 3 4 5

Strongly Disagree Strongly Agree

8. The people I interact with on Facebook would share things they value with me

1 2 3 4 5

Strongly Disagree Strongly Agree

9. I know people on Facebook well enough to ask them for favours.

1 2 3 4 5

Strongly Disagree Strongly Agree

10. The people I interact with on Facebook would help me fight an injustice.

1 2 3 4 5

Strongly Disagree Strongly Agree

11. Interacting with people on Facebook makes me interested in things that happen outside of my area

1 2 3 4 5

Strongly Disagree Strongly Agree

12. Interacting with people online/offline makes me want to try new things

1 2 3 4 5

Strongly Disagree Strongly Agree

Continued over page

13. Interacting with people on Facebook makes me interested in what people unlike me are thinking.

1 2 3 4 5

Strongly Disagree Strongly Agree

14. Talking with people on Facebook makes me curious about other places in the world.

1 2 3 4 5

Strongly Disagree Strongly Agree

15. Interacting with people on Facebook makes me feel like part of a larger community

1 2 3 4 5

Strongly Disagree Strongly Agree

16. Interacting with people on Facebook makes me feel connected to the bigger picture.

1 2 3 4 5

Strongly Disagree Strongly Agree

17. Interacting with people on Facebook reminds me that everyone in the world is connected.

1 2 3 4 5

Strongly Disagree Strongly Agree

18. I am willing to spend time to support general community activities on Facebook.

1 2 3 4 5

Strongly Disagree Strongly Agree

19. Interacting with people on Facebook gives me new people to talk to

1 2 3 4 5

Strongly Disagree Strongly Agree

20. On Facebook, I come in contact with new people all the time.

1 2 3 4 5

Strongly Disagree Strongly Agree

Appendix F Facebook Intensity Scales

Scale Items

1. Facebook is part of my everyday activity
2. I am proud to tell people I'm on Facebook
3. Facebook has become part of my daily routine
4. I feel out of touch when I haven't logged onto Facebook for a while
5. I feel I am part of the Facebook community
6. I would be sorry if Facebook shut down
7. Approximately how many TOTAL Facebook friends do you have? *

10 or less, 11–50, 51–100, 101–200, 201+

8. In the past week, on average, approximately how much time PER DAY have you spent actively using Facebook? **

= 0-14min, 2=15-29 min,

Response categories range from 1 = strongly disagree to 5 = strongly agree, unless otherwise noted.

*Can be asked as an open-ended (as in Ellison et al., 2007) or closed-ended (as in Steinfield et al., 2008) question. If asked as an open-ended question, Total Facebook friends must be transformed by taking the log before averaging across items to create the scale due to differing item scale ranges. If asked as a closed-ended question, a ten point ordinal scale may be used (e.g. 10 or less, 11–50, 51–100, 101–150, 151–200, 201–250, 251–300, 301–400, more than 400). You may wish to adjust these response categories depending on your population, etc. Note that earlier versions asked students to distinguish among in-network and total friends. This may or may not be appropriate based on population, site layout etc.

**Can be asked as an open-ended or closed-ended question. If asked as an open-ended question, Facebook minutes should be measured by having participants fill in the amount of time they spend on Facebook. Then the item should then be transformed by taking the log before averaging across items to create the scale due to differing item scale ranges. If asked as a close-ended question an ordinal scale may be used (e.g. 1= 0-14min, 2=15-29 min, etc). Again, response categories may differ based on population means.

Computing the Scale

The Facebook Intensity score is computed by calculating the mean of all of the items in the scale.

Taken from - <https://www.msu.edu/~nellison/TOIL/scales.html>

Appendix G Data Set 2 Facebook Status Updates Survey

The survey used for obtaining data for Facebook Status Updates can be viewed at

https://docs.google.com/forms/d/1OiopxZtLAGKiyAtQ9bOIn6VyvKDzOhB9Nd6S_N_411Y/viewform

Facebook Social Capital Survey

This is a survey intended to examine the social value that users obtain from Facebook. The first few questions are about your recent use of Facebook, the remaining questions are about the social benefit you might (or might not) be getting from the software.

The answers that you give will be treated confidentially. Your name will never be mentioned in the reporting of the results.

Participation is entirely voluntary and you don't have to answer all of the questions.

It will take approximately five minutes to complete the questions.

You might not benefit personally from the results but the summary results will be made available via <http://sociallearning.space.abertav.ac.uk/scism/>

In thanks for taking part in this survey you will be entered in to a prize draw for a £20.00 amazon gift voucher.

*Required

Agreement *

I agree to take part having read the explanation above. Please type in your facebook address (e.g. <https://www.facebook.com/d.carmichael.abertav>).

Facebook Use

It might be helpful to have your facebook page open to answer these questions - <https://www.facebook.com/>

How many friends do you have on Facebook?

How many status updates do you receive this last week?

How many comments did you receive this week?

How many likes did you receive this week?

How many status updates did you make last week?

How many comments did you make last week?

How many likes did you make this week?

Facebook Social Value

Please answer these questions with particular reference to how you have felt about the software over the last week.

1. There are several people on Facebook that I trust to help solve problems.

1 2 3 4 5

Strongly Disagree Strongly Agree

2. There is someone on Facebook I can turn to for advice about making very important decisions.

1 2 3 4 5

Strongly Disagree Strongly Agree

3. There is someone on Facebook that I feel comfortable talking to about personal problems.

1 2 3 4 5

Strongly Disagree Strongly Agree

4. If I felt lonely, there are several people on Facebook that I could talk to.

1 2 3 4 5

Strongly Disagree Strongly Agree

5. If I needed to borrow something of value, I know someone on Facebook I can turn to.

1 2 3 4 5

Strongly Disagree Strongly Agree

6. The people I interact with on Facebook would stick up for me.

1 2 3 4 5

Strongly Disagree Strongly Agree

7. The people I interact with on Facebook would speak well of me if asked to.

1 2 3 4 5

Strongly Disagree Strongly Agree

8. The people I interact with on Facebook would share things they value with me.

1 2 3 4 5

Strongly Disagree Strongly Agree

9. I know people on Facebook well enough to ask them for favours.

1 2 3 4 5

Strongly Disagree Strongly Agree

10. The people I interact with on Facebook would help me fight an injustice.

1 2 3 4 5

Strongly Disagree Strongly Agree

11. Interacting with people on Facebook makes me interested in things that happen outside of my area.

1 2 3 4 5

Strongly Disagree Strongly Agree

12. Interacting with people online/offline makes me want to try new things.

1 2 3 4 5

Strongly Disagree Strongly Agree

13. Interacting with people on Facebook makes me interested in what people unlike me are thinking.

1 2 3 4 5

Strongly Disagree Strongly Agree

14. Talking with people on Facebook makes me curious about other places in the world.

1 2 3 4 5

Strongly Disagree Strongly Agree

15. Interacting with people on Facebook makes me feel like part of a larger community.

1 2 3 4 5

Strongly Disagree Strongly Agree

16. Interacting with people on Facebook makes me feel connected to the bigger picture.

1 2 3 4 5

Strongly Disagree Strongly Agree

17. Interacting with people on Facebook reminds me that everyone in the world is connected.

1 2 3 4 5

Strongly Disagree Strongly Agree

18. I am willing to spend time to support general community activities on Facebook.

1 2 3 4 5

Strongly Disagree Strongly Agree

19. Interacting with people on Facebook gives me new people to talk to.

1 2 3 4 5

Strongly Disagree Strongly Agree

20. On Facebook, I come in contact with new people all the time.

1 2 3 4 5

Strongly Disagree ○ ○ ○ ○ ○ Strongly Agree

Thanks for completing the survey

!

Appendix H Data Preparation Data Set 2

The data set 2 was prepared by taking the survey results for the communication questions such as number of friends, comments, status updates, likes, posts and responses and verifying the data against the subject's timeline. The resultant numbers were mapped as follows:

SCiSM bonding = out-degree (outgoing status updates, comments and likes) multiplied by clustering coefficient (approximated*)

SCiSM bridging = actual in-degree (incoming comments and likes) divided by possible in-degree (friends)

* Clustering coefficient was approximated at .1 for all participants. The data was gathered from consenting participants in an open system. However ethical consent for data gathering of friend's names could not be obtained from a third party.

Appendix I Data Sets 1 & 2

Data set 1 - <http://sociallearningspace.abertay.ac.uk/scism/DataSet1.xlsx>

DATA FOR Participant 1 RP																												
Time	SCBond	SCBridg	SCISM	FIS	ISCS	ISCSBor	ISCSBridge	Correlations																				
								FIS	Corr	ISCS	ISCSBor	ISCSBridge																
1	0	0.1379	0.1379	3.8	39	16	23	SC	-0.181	0.8442	0.8084	0.7311																
2	0	0.1316	0.1316	3.9	39	19	20	SCBond	-0.195	0.8325	0.8096	0.7129																
3	0	0.1345	0.1345	3.9	39	18	21	SCBridg	-0.003	0.9222	0.7397	0.8325																
4	0.6667	0.1869	0.8536	3.9	53	24	29																					
5	0	0.1481	0.1481	4.4	43	18	25																					
6	0	0.1357	0.1357	4.1	43	20	23																					
7	0	0.137	0.137	4.0	42	19	23																					
8	0.1667	0.1636	0.3302	4.0	47	20	27																					
9	0	0.1357	0.1357	4.0	43	20	23																					
10	0	0.1368	0.1368	3.9	40	17	23																					
11	0	0.1299	0.1299	4.0	35	17	18																					
12	0	0.1307	0.1307	4.0	38	19	19																					
Means	0.0634	0.1424	0.2118	3.9792	41.75	18.917	22.833																					
FIS																												
ID	FIS 1	FIS 2	FIS 3	FIS 4	FIS 5	FIS 6	FIS 7	FIS 8	FIS Mean																			
1	4	3	4	4	4	4	5	4	3.8																			
2	4	3	4	4	4	4	5	4	3.9																			
3	4	3	4	4	4	4	5	4	3.9																			
4	4	3	4	5	4	5	4	2	3.9																			
5	5	3	5	5	5	5	4	3	4.4																			
6	4	3	4	5	5	5	4	3	4.1																			
7	4	3	4	4	5	5	4	3	4.0																			
8	4	3	4	4	4	5	4	4	4.0																			
9	5	3	4	4	4	5	4	3	4.0																			
10	4	3	4	4	4	5	4	3	3.9																			
11	4	3	4	5	4	5	4	3	4.0																			
12	4	3	4	5	4	5	4	3	4.0																			
ISCS																												
ID	Board 1	Board 2	Board 3	Board 4	Board 5	Board 6	Board 7	Board 8	Board 9	Board 10	Board 11	Board 12	Bridge	Bridge	Bridge	Bridge	Bridge	Bridge	Bridge	Bridge	Bridge	Bridge	Bridge	Bridge	Bridge	Bridge	Bridge	SC total
1	2	2	1	1	1	2	2	1	2	2	1	2	2	2	3	3	1	2	2	2	2	2	3	3	3	3	23	39
2	2	2	2	2	2	1	2	2	1	3	2	19	2	2	3	1	1	3	2	2	2	2	2	2	2	2	20	39
3	2	2	1	2	1	2	2	2	1	3	2	18	3	2	1	3	3	2	2	2	2	2	2	2	2	2	21	39
4	2	2	2	3	1	2	3	1	4	2	24	4	3	3	1	4	3	2	3	3	3	3	3	3	3	29	53	
5	2	2	2	2	1	2	1	1	2	3	18	2	3	2	1	3	2	3	3	3	3	3	3	3	3	25	43	
6	2	3	2	2	1	2	2	1	2	3	20	3	3	2	1	1	3	2	3	2	3	2	3	2	3	23	43	
7	3	3	2	2	1	2	1	1	2	2	19	2	3	2	1	3	2	2	3	2	3	2	3	2	3	23	42	
8	2	2	2	2	1	2	3	1	3	2	20	3	3	3	1	2	3	3	3	2	4	2	4	2	4	27	47	
9	3	2	1	3	2	2	2	1	2	2	20	2	3	2	1	2	3	2	3	2	3	2	3	2	3	23	43	
10	2	2	2	2	1	2	1	1	2	2	17	2	4	2	1	2	3	2	2	3	2	3	2	3	2	23	40	
11	2	1	1	2	1	2	2	1	2	3	17	2	3	2	1	2	1	2	1	2	2	2	1	18	35			
12	3	2	2	3	1	2	1	1	2	2	19	2	2	2	1	2	2	2	2	2	2	2	2	2	2	19	38	

DATA FOR Participant 2 FR																												
Time	SCBond	SCBridg	SC	FIS	ISCS	ISCSBor	ISCSBridge	Correlations																				
								FIS	Corr	ISCS	ISCSBor	ISCSBridge																
4	1	0	0.1379	0.1379	3.4	44	21	23	SC	-0.279	0.6844	0.595	0.6741															
5	2	1	0.1974	1.1974	3.5	56	26	30	SCBond	-0.282	0.6933	0.6044	0.681															
6	3	0	0.1345	0.1345	3.4	38	16	22	SCBridg	-0.238	0.5696	0.4754	0.5826															
7	4	0.6667	0.1869	0.8536	3.5	50	21	29																				
8	5	0	0.1481	0.1481	3.6	50	22	28																				
9	6	0	0.1357	0.1357	3.5	43	21	22																				
10	7	0	0.137	0.137	3.4	44	21	23																				
11	8	5	0.5841	5.5841	3.4	60	28	32																				
12	9	0.6667	0.181	0.8471	3.5	53	24	29																				
13	10	1.75	0.2051	1.3551	3.4	63	31	32																				
14	11	0.8	0.1818	0.3818	3.5	52	23	29																				
15	12	0.5	0.1634	0.6634	3.4	53	28	25																				
16	Means	0.8653	0.1934	1.0647	3.4479	50.5	23.5	27																				
FIS																												
ID	FIS 1	FIS 2	FIS 3	FIS 4	FIS 5	FIS 6	FIS 7	FIS 8	FIS Mean																			
1	5	2	5	3	4	3	3	2	3.4																			
2	5	2	5	3	4	3	3	3	3.5																			
3	5	2	5	3	4	3	3	2	3.4																			
4	5	2	5	3	5	3	3	2	3.5																			
5	5	2	5	3	5	3	3	3	3.6																			
6	5	2	5	3	5	3	3	2	3.5																			
7	5	2	5	3	4	3	3	2	3.4																			
8	5	2	5	3	4	3	3	2	3.4																			
9	5	2	4	3	5	3	3	3	3.5																			
10	5	2	5	3	4	3	3	2	3.4																			
11	5	2	5	3	4	3	3	3	3.5																			
12	5	2	5	3	4	3	3	2	3.4																			
ISCS																												
ID	Board 1	Board 2	Board 3	Board 4	Board 5	Board 6	Board 7	Board 8	Board 9	Board 10	Board 11	Board 12	Bridge	Bridge	Bridge	Bridge	Bridge	Bridge	Bridge	Bridge	Bridge	Bridge	Bridge	Bridge	Bridge	Bridge	Bridge	SC total
1	3	2	2	3	3	2	1	1	2	2	21	3	3	2	1	3	2	2	3	2	2	3	2	2	2	23	44	
2	4	2	2	3	1	3	3	2	3	3	26	3	3	2	1	4	4	3	3	4	3	3	4	3	30	56		
3	3	2	1	2	1	1	1	1	1	2	16	3	3	2	1	2	2	2	2	2	2	2	2	2	2	22	38	
4	3	2	2	2	1	2	3	1	2	3	21	4	5	4	1	2	2	2	2	3	3	3	3	3	29	50		
5	3	2	3	3	3	3	1	1	2	2	22	3	3	3	3	3	3	3	3	3	3	3	3	3	2	28	50	
6	3	3	2	2	1	3	1	1	2	3	21	2	2	3	1	2	2	2	3	2	2	3	2	2	3	22	43	
7	2	2	3	3	3	2	1	1	2	2	21	2	3	2	2	2	2	2	2	3	2	3	2	3	2	23	44	
8	3	2	4	3	2	2	3	3	3	3	28	3	3	2	2	4	3	4	3	4	3	4	4	4	32	60		
9	3	3	1	3	1	3	3	2	3	2	24	3	3	3	1	4	3	3	3	3	3	3	3	3	29	53		
10	3	3	3	4	3	3	2	3	4	3	31	3	3	4	2	1	4	4	4	4	4	4	3	3	32	63		
11	3	3	1	3	1	2	2	2	3	3	23	3	2	4	1	3	5	2	3	3	3	3	3	3	29	52		
12	3	3	1	3	3	3	3	3	3	3	28	3	2	1	3	2	2	2	2	4	3	2	2	4	3	25	53	

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
1	DATA FOR Participant 8 RF																										
2																											
3	Time	SCBond	SCBridg	SC	FIS	ISCS	ISCSBor	ISCSBridge																			
4	1	0	0.1379	0.1379	3.3	41	18	23																			
5	2	0	0.1316	0.1316	3.3	38	18	20																			
6	3	0	0.1345	0.1345	3.3	39	18	21																			
7	4	0.3333	0.1869	0.5202	3.3	47	22	25																			
8	5	0	0.1481	0.1481	3.3	47	22	25																			
9	6	0	0.1357	0.1357	2.3	42	19	23																			
10	7	0	0.137	0.137	2.4	37	19	16																			
11	8	0	0.1168	0.1168	2.5	36	19	17																			
12	9	0	0.0905	0.0905	2.3	34	18	16																			
13	10	0	0.1197	0.1197	2.3	34	18	16																			
14	11	0	0.1299	0.1299	2.3	36	19	17																			
15	12	0	0.1307	0.1307	2.4	41	20	21																			
16	Meanz	0.0278	0.1333	0.1611	2.7083	35.333	19.167	20.167																			
17																											
18																											
19																											
20																											
21	1	????	4	3	4	4	4	2	3	2	3.3																
22	2	0000165	4	3	4	4	4	2	3	2	3.3																
23	3	0000165	4	3	4	4	4	2	3	2	3.3																
24	4	0000165	4	3	4	4	4	2	3	2	3.3																
25	5	0000165	4	3	4	4	4	2	3	2	3.3																
26	6	0000165	2	3	2	2	2	2	3	2	2.3																
27	7	0000165	2	3	2	2	2	2	3	3	2.4																
28	8	0000165	2	3	2	2	2	2	3	3	2.5																
29	9	0000165	2	3	2	2	2	2	3	2	2.3																
30	10	0000165	2	3	2	2	2	2	3	2	2.3																
31	11	0000165	2	3	2	2	2	2	3	2	2.3																
32	12	0000165	2	3	2	2	2	2	3	3	2.4																
33																											
34																											
35	1	0000165	3	2	1	3	1	1	2	1	3	1	18	2	2	1	2	2	3	2	3	3	3	3	23	41	
36	2	0000165	3	2	1	3	1	1	2	1	3	1	18	1	2	3	1	3	2	3	1	1	3	20	38		
37	3	0000165	3	2	1	3	1	1	2	1	3	1	18	1	2	3	1	2	3	1	2	3	3	21	39		
38	4	0000165	4	2	1	3	1	1	4	2	3	1	22	2	3	3	3	2	2	2	2	2	4	25	47		
39	5	0000165	3	2	2	3	1	1	3	2	4	1	22	3	3	2	3	2	3	2	2	3	2	25	47		
40	6	0000165	3	2	1	3	1	1	3	1	3	1	19	2	1	3	3	1	3	2	2	4	2	23	42		
41	7	0000165	3	2	1	3	1	1	3	1	3	1	19	1	2	1	1	1	3	3	2	3	1	18	37		
42	8	0000165	3	2	1	3	1	1	2	1	4	1	19	2	1	2	1	2	2	3	2	1	1	17	36		
43	9	0000165	3	2	1	3	1	1	2	1	3	1	18	2	2	1	2	2	1	2	2	1	1	16	34		
44	10	0000165	3	2	1	3	1	1	2	1	3	1	18	1	1	2	2	2	1	1	2	2	2	16	34		
45	11	0000165	4	2	1	3	1	1	2	1	3	1	19	2	1	2	2	2	1	2	1	2	2	17	36		
46	12	0000165	4	2	1	3	1	1	2	1	4	1	20	4	2	3	2	2	2	1	2	1	2	21	41		
47																											
48																											

Data set 2 - <http://sociallearningspace.abertay.ac.uk/scism/DataSet2.xlsx>

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AI	
	1. There are several people on Facebook that have problems	2. There are several people on Facebook that have problems	3. There are several people on Facebook that have problems	4. If I had, there are several people on Facebook that have problems	5. If I had, there are several people on Facebook that have problems	6. The people I interact with on Facebook would like to talk to me	7. The people I interact with on Facebook would like to talk to me	8. The people I interact with on Facebook would like to talk to me	9. I know people who would like to talk to me	10. The people I interact with on Facebook would like to talk to me	11. I know people who would like to talk to me	12. I know people who would like to talk to me	13. I know people who would like to talk to me	14. I know people who would like to talk to me	15. I know people who would like to talk to me	16. I know people who would like to talk to me	17. I know people who would like to talk to me	18. I know people who would like to talk to me	19. I know people who would like to talk to me	20. On Facebook, I come in contact with people who talk to me	How many times did you use Facebook?	How many times did you use Facebook?	How many times did you use Facebook?	How many times did you use Facebook?	How many times did you use Facebook?	How many times did you use Facebook?	How many times did you use Facebook?	How many times did you use Facebook?	How many times did you use Facebook?	How many times did you use Facebook?
1	Timetamp																													
2	16/03/2015 10:37	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3	16/03/2015 14:02	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
4	23/03/2015 10:44	4	4	3	3	3	4	5	5	5	5	4	4	4	3	2	2	2	2	3	3	2	542	50	15	15	1	10	20	
5	23/03/2015 10:52	3	5	4	3	4	2	3	4	3	4	3	4	3	4	5	5	4	4	5	3	4	77	10	11	5	1	6	10	
6	23/03/2015 11:07	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
7	23/03/2015 11:08	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
8	23/03/2015 11:21	1	2	1	1	2	1	2	1	2	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
9	23/03/2015 11:24	1	4	2	2	2	4	2	4	2	4	2	3	4	4	4	5	4	5	4	5	4	5	4	5	4	5	4	5	
10	23/03/2015 11:31	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
11	23/03/2015 11:57	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
12	23/03/2015 12:15	4	3	5	4	4	4	5	3	4	4	4	2	2	2	2	2	2	2	2	2	2	3	925	40	24	11	1	20	20
13	23/03/2015 12:45	2	3	5	4	5	3	5	3	3	3	3	4	2	2	2	2	2	2	2	2	2	5	197	12	16	4	1	5	12
14	23/03/2015 13:00	3	4	3	3	4	4	4	4	4	4	4	2	2	2	2	2	2	2	2	2	2	2	1622	0	20	6	0	5	20
15	23/03/2015 14:03	1	2	1	1	2	1	2	1	2	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
16	23/03/2015 14:21	1	1	3	2	2	3	4	2	1	1	3	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
17	23/03/2015 14:51	3	2	3	2	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
18	23/03/2015 15:11	3	3	5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
19	23/03/2015 16:01	3	3	3	4	3	4	3	4	3	4	3	4	3	4	3	4	3	3	3	3	3	3	3	3	3	3	3	3	
20	23/03/2015 17:04	5	4	4	5	4	3	3	4	4	5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
21	23/03/2015 17:09	4	4	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
22	23/03/2015 17:36	3	4	3	3	3	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
23	24/03/2015 09:34	2	2	2	1	1	2	1	2	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
24	24/03/2015 09:46	2	1	1	2	2	1	2	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
25	24/03/2015 09:58	2	2	3	3	2	2	2	4	3	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
26	24/03/2015 10:33	4	4	4	5	4	5	3	4	4	5	4	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
27	24/03/2015 10:39	2	4	4	3	3	4	2	3	3	3	4	5	3	4	3	3	4	4	4	4	4	4	4	4	4	4	4	4	
28	24/03/2015 10:43	4	4	3	3	2	2	2	3	3	4	4	5	5	5	3	5	5	4	4	4	4	4	4	4	4	4	4	4	
29	24/03/2015 10:51	3	4	5	5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
30	24/03/2015 10:54	4	5	5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
31	24/03/2015 10:54	4	5	5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
32	24/03/2015 11:10	4	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
33	24/03/2015 11:19	2	2	3	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
34	24/03/2015 11:24	4	3	3	3	2	2	2	1	3	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
35	24/03/2015 11:49	3	3	3	2	2	2	4	4	4	3	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
36	24/03/2015 11:49	3	3	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
37	24/03/2015 12:10	2	1	1	1	1	1	2	3	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
38	24/03/2015 12:10	3	3	3	3	2	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
39	24/03/2015 12:44	2	2	4	4	5	4	5	4	5	4	3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
40	24/03/2015 12:44	2	2	4	4	5	4	5	4	5	4	3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
41	24/03/2015 12:44	2	2	4	4	5	4	5	4	5	4	3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
42	24/03/2015 12:44	2	2	4	4	5	4	5	4	5	4	3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
43	24/03/2015 12:44	2	2	4	4	5	4	5	4	5	4	3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
44	24/03/2015 12:44	2	2	4	4	5	4	5	4	5	4	3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
45	24/03/2015 12:44	2	2	4	4	5	4	5	4	5	4	3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
46	24/03/2015 12:44	2	2	4	4	5	4	5	4	5	4	3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
47	24/03/2015 12:44	2	2	4	4	5	4	5	4	5	4	3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
48	24/03/2015 12:44	2	2	4	4	5	4	5	4	5	4	3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
49	24/03/2015 12:44	2	2	4	4	5	4	5	4	5	4	3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
50	24/03/2015 12:44	2	2	4	4	5	4	5	4	5	4	3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
51	24/03/2015 12:44	2	2	4	4	5	4	5	4	5	4	3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
52	24/03/2015 12:44	2	2	4	4	5	4	5	4	5	4	3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
53	24/03/2015 12:44	2	2	4	4	5	4	5	4	5	4	3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
54	24/03/2015 12:44	2	2	4	4	5	4	5	4	5	4	3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
55	24/03/2015 12:44	2	2	4	4	5	4	5	4	5	4	3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
56	24/03/2015 12:44	2	2	4	4	5	4	5	4	5	4	3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
57	24/03/2015 12:44	2	2	4	4	5	4	5	4	5	4	3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
58	24/03/2015 12:44	2	2	4	4	5	4	5	4	5	4																			

Appendix J Correlations evaluation of the measures for all participants

SNA measures such as degree centrality, in-degree, out-degree, betweenness, closeness, eigenvector centrality and clustering coefficient where compared for significance against SCiSM, SCiSM bonding and SCiSM bridging. The text above contains a summary table from the following data.

Subject 1

	degree	in	out	Betweenne	Close	Eigen	Cluster	SCiSM	SCiSMBon	SCiSMBridge
ISCS	0.336448	0.104082	0.861294	0.172374	0.245059	0.461125	0.861294	0.84421	0.832497	0.922188315
ISCSBond	0.222439	-0.00377	0.796699	-0.18132	0.182047	0.342944	0.796699	0.808372	0.809566	0.73967192
ISCSBridg	0.356618	0.157796	0.764279	0.375842	0.246649	0.463865	0.764279	0.731149	0.712889	0.892457657
P-values	No	No	01	No	No	No	01	01	01	01
	No	No	01	No	No	No	01	01	01	01
	No	No	01	No	No	No	01	01	01	01

Subject 2

	degree	in	out	Betweenne	Close	Eigen	Cluster	SC	SCBond	SCBridge
ISCS	0.827473	0.781233	0.816578	0.53648	0.459341	0.785813	0.964198	0.776622	0.757953	0.673612903
ISCSBond	0.901966	0.804248	0.926784	0.609932	0.34767	0.660042	0.916485	0.884484	0.876352	0.489402725
ISCSBridg	0.661326	0.676334	0.612322	0.401135	0.535167	0.84385	0.918327	0.579012	0.550674	0.807265164
	01	01	01	No	No	01	01	01	01	05
	01	01	01	05	No	05	01	01	01	No
	05	05	05	No	No	01	01	05	No	01

Subject 3

	degree	in	out	Betweenne	Close	Eigen	Cluster	SC	SCBond	SCBridge
ISCS	0.634519	0.64569	0.619704	0.46664	-0.68599	0.073512	0.680886	0.684372	0.693254	0.56963195
ISCSBond	0.566715	0.588908	0.542076	0.391677	-0.68336	-0.02075	0.565498	0.594959	0.604422	0.475393201
ISCSBridg	0.608573	0.605998	0.606773	0.474862	-0.58104	0.164507	0.699421	0.674099	0.680954	0.582640464
	05	05	05	No	No	No	05	05	05	No
	No	05	No	No	No	No	No	05	05	No
	05	05	05	No	No	No	05	05	05	05

Subject 4

	degree	in	out	Betweene	Close	Eigen	Cluster	SC	SCBond	SCBridge
ISCS	0.841954	0.754335	0.890244	0.545863	-0.08467	-0.15278	0.944842	0.869622	0.865384	0.83553032
ISCSBond	0.870586	0.777348	0.923026	0.567619	-0.12756	-0.22889	0.926487	0.89694	0.895511	0.780186557
ISCSBridg	0.704926	0.636892	0.740294	0.450578	0.009852	0.015237	0.892973	0.732641	0.723134	0.868610486
	01	01	01	No	No	No	01	01	01	01
	01	01	01	No	No	No	01	01	01	01
	05	05	01	No	No	No	01	01	01	01

Subject 5

	degree	in	out	Betweene	Close	Eigen	Cluster	SC	SCBond	SCBridge
ISCS	0.871718	0.802415	0.820477	0.737221	-0.3438	0.429131	0.952744	0.857773	0.799352	0.680867269
ISCSBond	0.891938	0.785306	0.899998	0.593543	-0.40072	0.227242	0.8047	0.910603	0.897057	0.406456269
ISCSBridg	0.760799	0.726759	0.671293	0.762449	-0.26381	0.531366	0.957481	0.724246	0.639031	0.809086822
	01	01	01	01	No	No	01	01	01	05
	01	01	01	05	No	No	01	01	01	No
	01	01	05	01	No	No	01	01	05	01

Subject 6

	degree	in	out	Betweene	Close	Eigen	Cluster	SC	SCBond	SCBridge
ISCS	0.911369	0.838168	0.92781	0.703039	0.170999	0.524757	0.911313	0.877578	0.86826	0.899418138
ISCSBond	0.948756	0.895904	0.946569	0.703547	0.041521	0.434072	0.872507	0.896065	0.88749	0.90061135
ISCSBridg	0.815086	0.729518	0.846406	0.653156	0.270429	0.565911	0.880622	0.79994	0.790638	0.835131849
	01	01	01	05	No	No	01	01	01	01
	01	01	01	05	No	No	01	01	01	01
	01	01	01	05	No	No	01	01	01	01

Subject 7

	degree	in	out	Betweene	Close	Eigen	Cluster	SC	SCBond	SCBridge
ISCS	0.802656	0.780752	0.818862	0.401664	-0.05884	0.766245	0.796588	0.835431	0.831645	0.851393503
ISCSBond	0.785447	0.763157	0.802212	0.368565	-0.12352	0.722635	0.807727	0.844248	0.842411	0.823011504
ISCSBridg	0.778386	0.758284	0.792896	0.422115	0.030709	0.779258	0.734942	0.774593	0.768434	0.839132444
	01	01	01	No	No	01	01	01	01	01
	01	01	01	No	No	01	01	01	01	01
	01	01	01	No	No	01	01	01	01	01

Subject 8

	degree	in	out	Betweene	Close	Eigen	Cluster	SC	SCBond	SCBridge
ISCS	-0.25688	-0.36608	0.544007	0.462596	0.557492	0.811116	0.544007	0.619876	0.544007	0.824190872
ISCSBond	0.070081	-0.03956	0.608308	0.596598	0.2228	0.489692	0.608308	0.651949	0.608308	0.710524235
ISCSBridg	-0.37068	-0.46726	0.453957	0.35132	0.640452	0.859402	0.453957	0.535289	0.453957	0.780102903
	No	No	No	No	No	01	No	05	No	01
	No	No	05	05	No	No	05	05	05	01
	No	No	No	No	05	01	No	No	No	01