A Novel Application of Cognitive Agent Based Computing Framework for Influence of Choices in Social Scenarios

By

Faiza Ghafoor

Department of CS/SE

Faculty of Basic and Applied Sciences

International Islamic University Islamabad

2015

1. Intelligent agents (Computer software)
2. computer Science

Department of Computer Science and Software Engineering International Islamic University Islamabad

Date: 11-Dec-2015

It is certified that we have read the thesis entitled as A Novel Application of Cognitive Agent Based Computing Framework for Influence of Choices in Social Scenarios submitted by Miss. Faiza Ghafoor Reg # 318-FBAS/MSSE/F10 and it is our decree that this dissertation is of satisfactory standard to certify its acceptance by the International Islamic University Islamabad, for MS degree in Software Engineering.

COMMITTEE

External Examiner

Dr. Malik Najm us Saqib
Assistant Professor
Department of Computer Science
CIIT Islamabad

Najm

Internal Examiner

Miss Saima Imtiaz
Lecturer
Department of CS & SE
IIU Islamabad

Supervisor

Dr. Muaz A. Niazi

Professor
Department of Computer Science
CIIT Islamabad

Co Supervisor

Mr. Asim Munir

Assiatant Professor
Department of CS & SE
IIU Islamabad

Acting Chairperson

Miss Asma Batool

Lecturer
Department of CS & SE

IIU Islamabad

Department of Computer Science and Software Engineering International Islamic University Islamabad

Date: 11-Dec-2015

It is certified that we have read the thesis entitled as A Novel Application of Cognitive Agent Based Computing Framework for Influence of Choices in Social Scenarios submitted by Miss. Faiza Ghafoor Reg # 318-FBAS/MSSE/F10 and it is our decree that this dissertation is of satisfactory standard to certify its acceptance by the International Islamic University Islamabad, for MS degree in Software Engineering.

COMMITTEE

External Examiner

Dr. Malik Najm us Saqib
Assistant Professor
Department of Computer Science
CIIT Islamabad

Internal Examiner

Miss Saima Imtiaz
Lecturer
Department of CS & SE
IIU Islamabad

Supervisor

Dr. Muaz A. Niazi

Professor
Department of Computer Science
CIIT Islamabad

Co Supervisor

Mr. Asim Munir

Assiatant Professor
Department of CS & SE
IIU Islamabad

Acting Chairperson

Miss Asma Batool

Lecturer
Department of CS & SE

IIU Islamabad

بسم الله الرحمن الرحيم

Abstract

Modern software Engineering requires handling large scale data, for example in the form of graphs or complex networks. Complex network studies mainly focus on the study of real world networks such as computer or social network. In the past, frameworks such as cognitive agent based computing and formal agent based simulation have been introduced to study the behavior of complex systems however cognitive agent-based computing has previously never been used to study the influence of social choices.

In this thesis I present a novel application of cognitive agent-based computing framework for modeling human behavior. The main goal is to develop a hybrid model made of formal specification model coupled with complex network modeling which allows us to analyze human behavior. In this study, structural characteristics of the complex network are analyzed and the centrality measures are calculated to identify actor's prominence and their behavioral influences in a social network. Statistical analysis is also performed to analyze the behavioral influence of girls over their friends located at the first degree of separation.This methodology presents an idea of modeling human behavior with the help of developed hybrid model in order to get better understanding of human behavior. This will be helpful in many software engineering practices such as requirement elicitation for large-scale networks where targeting single individual for requirement elicitation seems impossible.

Dedication

Dedicated to my Parents for their prayers, kindness and patience.

Acknowledgment

Firstly, I would like to express my endless gratitude to Allah Almighty (S.W.T) for His abundant blessings. Without His guidance, affection and love it was really impossible for me to accomplish anything in my life. He is the creator who helps me in the dark hours of my life. He is the only one who creates the way when there seems to be no way. He comforts me when I feel agonized, He guides me when confusion overtakes me.

I owe my gratitude to my supervisor Prof. Muaz Niazi who believed in me and gave me the chance to pursue the goal. He supported me through the thesis in a best possible way. I received his support during several hard periods and I would like to thank him for his continuous efforts to show me the path to follow. It was an honor for me to work with him.

I am also grateful to my Co-Supervisor Mr. Asim Muneer who supported me in every step.

I would also like to show my heartiest gratitude to my brother Mr. Ali Noor for his sincere and kind support. He always advised me in the best way and show me the way to get through the hardest time.

I would also like to thank my friends and siblings for their love and care. They impelled me to believe that true relations are indeed a great blessing of Allah. I am grateful to them for always listening to me, encouraging me, and making me laugh even when I didn't manage to smile.

Declaration

I hereby declare that the research presented in this thesis is my own work, excluding where otherwise acknowledged and that the thesis is my own composition. No part of the thesis has been previously presented for any other degree.

Date: 11 Dec 2015

Full Name: Faiza Ghafoor

Contents

List of Figures

List of Tables

1 Introduction

Modern software Engineering requires handling large-scale data, for example in the form of graphs or complex networks. Complex network studies mainly focus on the study of real-world networks such as computer networks or social networks. Network analysis is an important method for collecting, accumulating, storing, modeling and visualization of large-scale data obtained from the different complex adaptive system. In complex adaptive systems, interactions can be modeled as networks. The key difficulty in network modeling is to understand which data needs to be captured to develop the network model. Once the data has been converted to a network, network analysis techniques, mathematical tools and software tools can be used for the analysis [53]. In the current era, online and offline big data sources are easily available. However what is missing is the extensive application of computational methods to discover hidden patterns or emergent patterns in the data. A better understanding of these patterns can be used for the betterment of the society and humanity. Emergent behavior can be considered as a means of explaining collective behavior or properties of groups of components or agents or else it can also be used to describe a system in relation to the environment [50].

In the past, frameworks such as cognitive agent-based computing and formal agent-based simulation have been introduced to study the behavior of complex systems however cognitive agent-based computing has previously never been used to study the influence of social choices. Niazi and Hussain have introduced this framework for the modeling and simulation of complex adaptive system (CAS), this framework is called cognitive agent based computing (CABC) and is used for developing a deeper understanding for complex systems. The main goal of this CABC framework is to help in developing computational models by using Agent-based (where agents are modeled in a simulation) or complex network based methods (where nodes are connected in the form of the graph or network) [50].

Hostel life is a part of a complex social system. It emerges as a result of non-linear and complex interactions of various agents. Hostel residents are social agents in this complex system. These social agents have complex interactions with each other. Students also have much diversity in terms of their family backgrounds, attitudes, behavior and personal choices. Understanding the hidden dynamics of these systems can lead not only to an understanding of social behavior but also to a deeper understanding of the entire society. To study this complex social system, we have applied a complex network modeling level of CABC framework by using social network analysis. The key feature of the social networks is the users and their connections [39]. These connections made the structure of the network. Structural characteristics have been enormously used to understand the human behavior in social networks [67]. Social networks can be analyzed by analyzing structural characteristics of the network. Several measures have been suggested yet to analyze the structural characteristic of the network. One of these is centrality measure. The centrality measures are important because they tell us about the most influential and powerful node in the network [17]. Many centrality measures have been proposed yet; the most famous ones are degree centrality, closeness centrality, eccentricity and betweenness centrality [34]. The fifth, most popular one is Eigenvector centrality measure which has been proposed by Bonacich [13]. This centrality measure is usually used to find the most influential person in the network. There are numerous studies have been done on the social influences and behaviors of people [18] [26] [5].

1.1 Problem Statement

In the past, many frameworks have been introduced to study the behavior of complex systems however cognitive agent-based computing has previously never been used to study the human behavior in order to understand the complex interactions of humans for the future implication of software practices.

1.1.1 Original Contribution

Modeling framework we have used here involves two phases of model construction.

1. Modeling framework of CAS using a mathematical framework originally based on a formal specification language Z in order to represent the clear and unambiguous pattern of choices and behavior of people.

2. Complex network based model is developed as a validation of this study.

1.2 Overview of Thesis

The main contribution of the thesis is the hybrid model of behavior using Z- specifications and complex network modeling. Then two case studies are conducted.

First is based on statistical analysis. Behavioral patterns of actors participating in a complex social system are discovered.

The second case study is based on a "Complex network modeling level of CABC framework". After modeling of the complex network, centrality based analysis is conducted and then we have also analyzed the behavioral patterns of the first degree of friends of a focal node participating in offline complex social networks.

Thesis Outline Chapter 2: This chapter presents background and related work for the better understanding of the thesis.

Chapter 3: This chapter presents the formal specifications based on Z language.

Chapter 4: This chapter presents the methodology of research in conjunction with two different case studies. They were conducted on two different data sets collected from the hostels of International Islamic University Islamabad.

Chapter 5: This chapter presents detailed results of conducted social network analysis and also the results contribution of case studies.

Chapter 6: This chapter concludes the whole thesis.

2 Background

This chapter provides the background knowledge for the thesis.

Section 2.2 gives the basic concepts of social networks, network types, different network measures in which we specifically focused on centrality measures in our research. Then we have provided previous methods and finding on behavioral influences in large social networks.

Section 2.3 gives an overview of complex systems, the concept of emergence and the idea of CAS. Then we further discussed the different types of CAS modeling and finally modeling framework and their implementations are discussed to link the complex system studies with the diffusion of behavior in large social networks.

First of all, we have presented our selected reviews and their details.

Journal Name	Year of publication	Impact Factor	No. of Papers
The journal of Learning Science	2006	3.26	1
Elsevier (Social Network)	2012,2010,2006,1995,1979	3.137	7
Elsevier (Addictive Behavior)	2014, 2007, 2006	2.441	3
Springer	2012	1.4	1
Psychological Science	2014	4.43	1
PlosOne	2013,2010	3.534	5
Journal of Mathematical Sociology	1972	1.04	1
American Journal of Sociology	1991	3.476	1
New England Journal	2008,2007	55	1
British Medical Journal	2009	17.44	1
Proceedings of Royal Society	2012	1.971	1
Annals of Internal Medicines	2010	17.810	1
PNAS	2014	9.809	1
Association for Psychological Science	2014	4.543	1
IEEE	2011	1.438	
Journal of Computer Mediated Communication	2007	3.117	1
Nature	2001	36.45	1
SIAM	2003	1.175	1

Figure 1: Selected Studies for review

2.1 Complex System

Bar-Yam stated that,

A complex system is a system made up of many interacting components that shows their emergent behavior, which can't be understood from the behavior of its component alone [7].

Many complex phenomena are focused by different fields of science.

Complex systems can be distinguished from the complicated systems as in complex systems if any small part of the system is removed, the whole property of the system gets changed.

Emergence Bar-Yam [8] defines emergence as,

Emergence is the arrangement of collective behaviors of the interacting systems that they would not do alone.

Uri Wilensky [40] explained the phenomenon of the emergence of patterns, he stated that there are so many patterns in the world. Large scale patterns come from the interactions of the numerous small number of components. These components combine in an interesting way and create large-scale patterns. He called them as "macro patterns" that arise from the combination of "micro-agents". This phenomenon is called the emergence of patterns. Birds in the flock flying in the sky are the simplest example of emergent patterns. They don't have any leader birds to follow; their pattern emerges out from their behavior and adjustment in interaction with each other.

2.2 Complex Adaptive System

Mitchell [48] stated that:

Complex Adaptive Systems (CAS) are assumed to be the subset of complex systems which form from the non-linear interaction of the interacting units.

CAS are studied as a subset of the complex system. The dynamically changing environments of components and their simple self-organized local relationships lead to the emergence. The regularity in the emergent behaviors of the system is known as the pattern of the emergent behavior of a specific complex system [4].

In general CAS is considered as a special form of natural and artificial complex systems "where whole is more than the parts" and the observed complex behavior of the specific CAS is known as the "emergent behavioral pattern" of that system. For example, human body is a CAS, cells interaction at the micro level in the human body doesn't tell us about the actual capabilities of the human body i.e, eating, moving, sleeping, reading etc.

2.2.1 Modeling Complex Adaptive System

To develop the deeper understanding of a CAS, it is required to develop special types of models called the explicit models [25] and implicit models [3].

Implicit models work for mental cognition whereas explicit models are suitable for the communication purposes among researchers. However, the goal of the explicit model is to improve implicit model which leads to the better understanding of the CAS.

Here we are discussing two modeling approaches preferred by the CAS researchers.

- Agent Based [6].
- Complex Network Based [49].

Agent-Based Modeling The one who acts can be called as agent [50]. This modeling approach is used to observe simple simulated agents interacting with a large population of other agents to identify their global behaviors. Netlogo is a commonly used tool for this modeling approach.

Complex Network Based Modeling Complex network-based approach is originated from the graph theory; where complex interactions of different actors (also termed as nodes) are modeled in a form of complex network.

Each node (actor) and edge (relation) are loaded with information in a software tool to analyze the emergent behaviors.

Many tools are used for complex network modeling such as Pajek, UCINet, Gephi etc.

Structure of the Complex Networks To analyze the structure of the network by characterizing, describing and extraction of information has become one of the main goals of science. This study of a network has drawn attention by many fields of research such as computer science, biology, social science, economics and many more.

Two approaches have been used for the analysis of the complex network, one based on "topological structure" and the other is based on "dynamic properties" of a network.

The topological approach helps in understanding the functionality of the system [4]. Strogatz [71] has noted in his study that "structure effects function". Giorani Scardoni and Carlo Laudanna [65] have explained the examples of structural effects over functionality as the topology of the road network structure causes a critical traffic jam. The topological structure of social network effects over information diffusion. To understand the topology of a network, many principles have been studied in different fields of science such as cluster [49], network motifs [49], scale-free network [76], small world, centrality etc.

2.2.2 Modeling Frameworks

Niazi and Hussain have introduced different frameworks for the modeling and simulation of the complex adaptive system (CAS).

Cognitive Agent Based Computing Framework (CABC) This framework is developed to get a deeper understanding for the different complex adaptive systems.

The first phase of "Cognitive Agent Based Computing" refers to "cognition" of different phases of CAS under study. The next part of this framework is "Agent based computing" refers that in each CAS actors (nodes) and their relations (edges) are focused.

The objective of this framework is to develop computational models by using agent or complex network based methods [50].

This framework has different levels.

- First one is the Complex Network Modeling level.
- Second is the Exploratory Agent-Based Modeling level.
- Third is Descriptive Agent-Based Model Approach (DREAM) [51].
- The Fourth one is Virtual Overlay Multi-Agent System (VOMAS) [53] approach.

Different researchers choose the level of the framework according to their research goal.

Formal Agent Based Simulation Framework (FABS) An established modeling framework is developed [70] to effectively model and analyze a different type of CAS.

FABS modeling framework involves two phases of model construction,

- The First level is the formal model of CAS using a mathematical framework of formal specifications written in Z.

- The second level is Agent-Based Model which is based on formal specifications and is developed as a proof of concept.

Framework for the phenomenon of the disease spread A formal modeling framework using Z was proposed in conjunction with agent-based simulation model to understand the complex phenomenon of disease spread. They have used formal specification model for HIV/AIDS spread using Z-specification language for the clear representation of the complex social network of people suffering from HIV. Then they translated these specifications into ABM which is a first level of CABC framework [52]. This study has been done as an example of the complex adaptive system.

2.3 Social Network

Wasserman and Faust [75] have defined social network as:

A social network is a collection of people who have relation or relations between them.

2.3.1 Fundamentals of Social Network

Social network scholars are trying to know about complex relations between the nodes, by analyzing the topological structures of the networks [57]. The basic concepts of network analysis are discussed below.

Actor An actor acts as a building block of the network. The social entity whose presence makes the structure of the network is known as an actor. In a graph theory, this is known as vertex [75]. In computational sciences, this is known as agent [50]. For example individual in a group, a cell in an embryo, bird in a flock etc.

Edge Edges are the representation of relational ties of the actors. This shows how actors are connected with each other. By giving weight to the edges, the strength of the tie, emotional intensity, exchange of services, function of duration can be represented [30].

Dyad Dyad is a pair of an actor who have tie between them. The dyadic analysis used to analyze pairwise relationship such as husband wife [73], male-female partnership [38] etc.

Triad A relationship between three connected actors is called triad. Triadic analysis depends upon the social system where we need to analyze triangular relationships [75].

Subgroup A Collection of actors more than 3 and the associated ties among them is known as subgroup. It could also be a collection of dyads and triads [75].

Group The Group is a collection of all actors whose relations / ties can be measured [75].

Relation A Relation consists of different types of ties bounded in a specific scenario [75].

Social Network Social network is a collection of people who have relation or relations between them [75].

Social Network Analysis Wasserman and faust states that:

The idea of social network includes theories, models and applications that are expressed in terms of relational concepts or processes. Social Network Analysis (SNA) is based on the idea of ties among interacting units [75].

Centrality measures are the most popular and effective measure to know about the importance or prominence of the actor.

2.3.2 Types of Network

There are different types of social networks that are studied by the researchers. The nature of the group of actors and their relational properties determine the type of the network.

Most commonly discussed network types are:

One Mode Network In one mode network, a single set of actors is analyzed.

Two Mode Network In two-mode network, two different sets of nodes are analyzed. For example events and people etc.

Ego-Centered Network and Dyadic Network Ego-Centered network consists of a focal node termed as "Ego" and a set of "alters" who have the relation with the ego.

For example: In a mother and kids network; mother is a focal node and the kids have ties to the mother. In a dyadic network, the relationship between two actors is measured i.e, husband-wife, father-son etc.

2.3.3 History of Centrality Measures

The concept of centrality was introduced by Camille Jordan in 19th century [31].

To identify the most central node than others is a key issue in network analysis [15] [14].

In 1948, Bevalas has introduced the idea of centrality. He tested the hypothesis of the relationship between centrality and influence in groups.

Cohn and Marriott (1958) used this concept of centrality for understanding Indian political integration.

Pitts (1965) used this concept to understand the river transportation network in Russia and the preeminence of Moscow city.

In 1974, Czepiel used centrality measures to identify the patterns of diffusion [27].

In 1978; Freeman had formalized the degree; closeness and betweenness centrality to find the positions of the node whether a focal node has how many ties, how quickly it can reach all other nodes and how efficiently it can control the flow of information between other nodes respectively. However about closeness it was said that, it does not work over the disconnected network, it only works over the largest component of a network [57].

In 1995, Frank Harary and Prehage has introduced a new centrality model known as eccentricity. He illustrated this centrality measure as a path center of a graph and applied it on an example of two islands network in Oceania. He stated eccentricity as "e" of node v, a maximum distance "d" for all "u" nodes in a connected graph "G". Eccentricity and closeness were proved as two closely related centrality measures as eccentricity refers to the maximum distance a node has to any other node.

$$Ecc(v) = \frac{1}{Max\{(dist(v,t))\}} \tag{2.1}$$

Whereas as closeness refers as the total distances from a node to all other nodes in a network.

$$Close(v) = \frac{1}{distsum(v)} \tag{2.2}$$

In 1972, Bonacich [13] suggested eigenvector centrality; this centrality measure works on the weighted networks, unlike degree. This centrality not only works as a function of its neighbors but also a function of the number of neighbors of neighboring nodes [2]. So this centrality measures consider the whole pattern of the network.

2.3.4 Centrality Measures

Degree Centrality A number of links a node has presents the degree centrality of a node. Directed networks are analyzed by using the metrics of indegree and outdegree. Indegree is referred as the number of links coming towards nodes. Regarding our friendship network, indegree correlates with the popularity of girl among her friends. Outdegree is referred as the number of links going out from the node. In the friendship network outdegree correlates with the social nature of the girl in a network.

Degree Centrality was first formulated by Freeman [27]. It can be calculated by using Freeman's formula of Degree centrality:

$$C_D(v) = \frac{K_v}{n-1} = \sum\nolimits_{j \in G} \frac{a_{vj}}{n-1} \tag{2.3}$$

Where Kv is the number of connections of a node and n is the total number of nodes in the network.

Betweenness Centrality According to the Freeman's approach betweenness centrality is defined as the proportion of times that the node acts as a bridge between two different nodes for sending information. So this is the most favored position of the actors in the network. So the actor offering most shortest (geodesic) pathway between other pairs of actors tend to have high betweenness centrality [35].
Betweenness centrality can be calculated using the Freeman's formula:

$$C_B(v) = \sum_{s \neq v \neq t} \frac{\delta st(v)}{\delta st} \tag{2.4}$$

where δst represents sum of the shortest path, where s and t can be considered as two different nodes. $\delta st(v)$ represents sum of paths that intersect node v.

Closeness Centrality Closeness Centrality referred as a how far a node from all other nodes in the network [35]. If we consider the distance of node as a farness of the node. The closeness would be "the inverse of the farness". Less total distance of a node, lesser will be the closeness centrality. A node having low closeness centrality considered to be more central in the network. Closeness centrality can be calculated using the Freeman's formula:

$$C_C(v) = \sum \frac{1}{dist(v,t)} \tag{2.5}$$

Here v and t are the nodes.

Eccentricity Centrality Eccentricity centrality is quite close to the closeness centrality measure. It can be calculated by finding the smallest path of a node to all other nodes in the network, then the "longest" smallest path is considered between the two most distant nodes. Once the later one is identified, its reciprocal is calculated. The node having low eccentricity is considered to be the most proximal node in the network and so as easily reachable by others [65] and can be more influential than others.
On the contrary, the node having higher eccentricity shows the lower rate of information diffusion due to high structural proximity [10].
It can be calculated as :

$$C_{Ecc}(v) = \frac{1}{Max\{(dist(v,t))\}} \tag{2.6}$$

Here "v" and "t" are the nodes.

Eigenvector Centrality Bonacich's Eigenvector Centrality is used to calculate the centrality of a node as a function of the centrality of its neighbors. Eigenvector centrality focuses on the fact that a node having the connection to high centrality nodes are more important than the node having links with low centrality nodes [16]. This is a measure for finding the influential node in a network.

$$\lambda v = Av \tag{2.7}$$

λ is constant, v is the eigenvector and A is the adjacency matrix.

2.3.5 Correlation of Centralities

Rothenberg et.al [63] analyzed relationship between centrality measures and concluded that degree and betweenness are highly correlated measures.

Wasserman et.al [75] stated that degree, betweenness, eigenvector and closeness are all used to find the prominent actors in the network. There exist some conceptual over lap between these measures and also they are conceptually distinct to each other though this distinction depends on the topology of the network.

Friedkin [28] stated that closeness measure is related to the idea of independence and efficiency as the actor having higher closeness centrality does not need to rely on other actors for the information and can transmit information efficiently due to closer position than others. About betweenness, he stated that the actor having high betweenness have the ability to influence nearby actors in the network and can perform a great role for the flow of information through the network. The actor having high betweenness have the potential to facilitate, hinder or even alter the communication between neighboring actors [27, 49] .

Niazi and Batool [52] stated that closeness and eccentricity are conceptually similar to each other, the quality which distinct them is that closeness utilizes minimum distance whereas eccentricity calculates maximum geodesic distance between the focal actor and all other actors and they also found correlation between closeness and eccentricity measure. They also stated that degree centrality also correlates with eigenvector centrality whereas for betweenness they concluded that it variate according to the topology of the network.

2.3.6 Social Network and the Spread of Behavior

In the past, social network analysis have been used enormously to analyze the behavioral influence and the spread of habits or information across the network. Christakis et.al have widely presented their research over the spread of behavior across the networks by using social network analysis and statistical analysis. Specifically he have observed the spread of obesity [17], happiness in human society [26], emotional states either positive or negative [37], social contagion of sleep behavior and drug use behavior [47], spread of smoking behavior [18], alcohol consumption [62] etc. In all these studies they have took a large social network of people and applied centrality measures and statistical analysis such as logistic regression and generalized estimation equation (GEE) procedures on longitudinal data. In all these studies they have focused on the spread of behavior.

To analyze the spread of academic success, Hiroki Sayama [12] has also used social network analysis, linear regression and correlation and observed the relatively higher influence of intermediate level friendship for the spread of academic success in a school students. To overcome the spread of disease in Racoon's population [38], social network analysis have been used to analyze the meeting pattern and transmission rate of the disease. To understand the global influence of language such as German, French, Spanish, Chinese, Arabic, Hindi and English. Researchers [59] have mapped the networks of book translations, wikipedia specialists and network of people who tweets in two different languages. By using eigenvector centrality and correlation, they found that English is the central hub and most dominating language in the network.

In some other studies, researchers focused over the suicidal behavior [46], country and gender difference influences on drinking behavior [5], friends influence on future smoking behavior [33], HIV spread through needle sharing [74], quality of spouse/partner relationship influence on one's blood pressure [73].

Hill et.al [37] have also suggested a theoretical framework (SISa model) for studying the phenomenon of transmission of behavior and emotional health states.

In all these studies, they have analyzed the spread of behavior in social networks by using social network analysis in conjunction with statistical analysis to get interesting patterns of behavior of people in a large social network.

3 Methodology

The purpose of this study is to develop a hybrid model in order to understand human complex interactions as an example of complex system studies. This can be beneficial for the future social network sites as well as many other software engineering practices.

The main objective of this chapter is to (1) describe the research methodology of this study (2) introducing the Z based formal specifications adopted from FABS framework (3) introduce the complex network modeling level 1 adopted from CABC framework (4) describe both case studies conducted as a proof of concept of Z based specifications.

Additionally we have also explained the procedures used for the development of complex network model and also described the network analysis and statistical procedures to analyze the data.

3.1 Overview

Our methodology starts with the formal modeling of complex system using a mathematical framework originally based on a formal specification language Z. Secondly we have conducted case study using statistical analysis to analyze the behavior. Thirdly, complex network modeling level 1 of CABC framework is applied to conduct another case study. This case study is developed as a validation study based on formal specifications in order to study human behaviors.

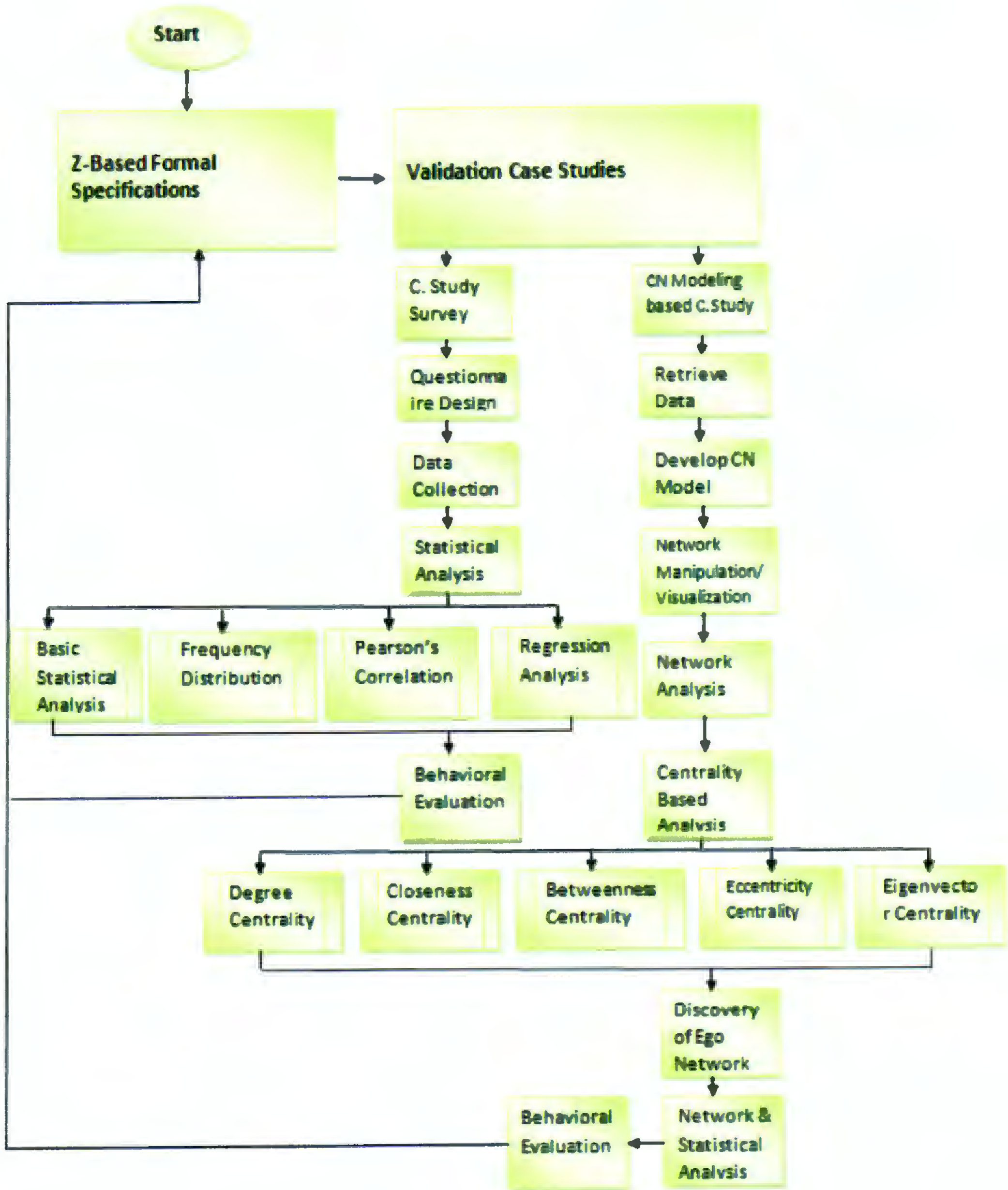


Figure 2: Research Methodology Diagram

3.2 Formal Specifications

Formal specifications are discussed in detail in the next chapter

3.3 Complex Network Modeling

The complex network modeling is adapted from the "Cognitive Agent Based Computing Framework" [50]. This framework has four levels. First one is the Complex Network Modeling level, second is the Exploratory Agent-Based Modeling level, third is Descriptive Agent-Based Model Approach (DREAM) [51] and the fourth one is Virtual Overlay Multi-Agent System (VOMAS) [53] approach. Researchers adopt the level which suits their research goal.

We adopted the first level "Complex Network Modeling" from this framework for our research. We have explained this framework level in our second case study by applying on our interaction data.

3.4 Validation Case Studies

We have conducted validation case studies. First case study is based on statistical analysis to analyze the behavior of the girls. Second case study is the application of cognitive agent based computing framework in which complex network modeling level of CABC framework is used to evaluate the behaviors of people participating in a complex network study.

3.4.1 Case Study 1

This case study is based on the statistical analysis in order to analyze the specific behaviors and choices of people.

Data Collection We obtained all the data from the girls hostel of International Islamic University Islamabad by developing the questionnaire related to their choices of dresses. Most of them were related to their choices of dresses. We got 50 responses ($responserate = 100\%$) that were included in statistical analysis ($n = 50$).

Statistical Analysis We performed statistical analysis on the data to understand the emergent patterns of the choices of girls. We found an interesting pattern that showed how girl's age correlates with their inclination towards dressing choices.

Basic statistical analysis involves developing statistical summaries of the attributes of the people and frequency distribution of the variables. We also applied Pearson's product moment correlation test among different variables with the alternative hypothesis that the true correlation is less than 0. We calculated 95 confidence intervals for each variable. Single sample T-test on age was applied by hypothesizing that average age is greater than 20. To determine how age affects choices about dresses. We conducted Pearson's correlation coefficient and found the extent of an interpersonal relationship of age and their dressing interests. We also performed linear regression analysis and evaluated the extent of dependency of different variables by taking age as an independent variable.

3.4.2 Case Study 2

This case study is based on Complex network modeling level 1 of CABC framework. The complex network modeling is applicable when we have suitable interaction data.

Our data set has 498 nodes (people). First of all, we made a graph in the form of the complex network by using interaction data and then perform complex network analysis.

This case study is performed by the combination of different tools namely Gephi[9], Pajek[20] and R[43]. Gephi helped us in performing visualization-based analysis of the girls living in the hostel. Pajek helped us in performing detailed network analysis and R is used for simple statistical analysis.

Development of Complex Network Model Five steps have been performed for the development of complex network model.

1. Retrieve Interaction Data
2. Develop Complex Network Model
3. Manipulate Complex Network
4. Perform Complex Network Analysis
5. Discover Emergent Patterns

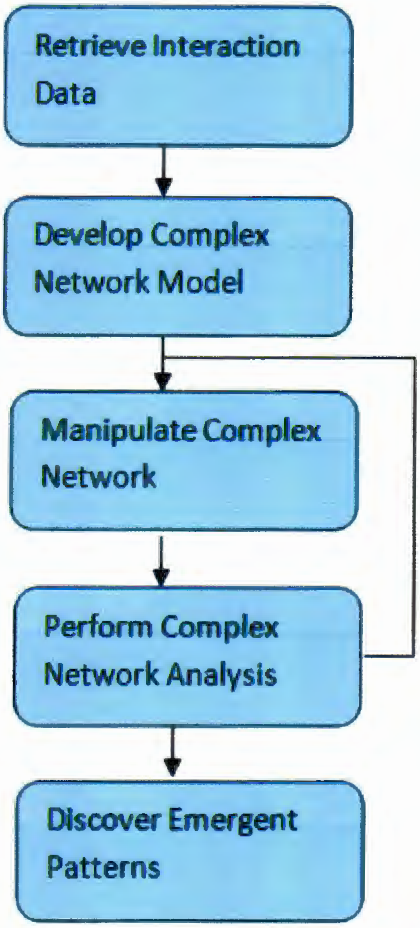


Figure 3: Complex Network Modeling adapted from CABC framework

Retrieve Interaction Data There are many ways in which network data can be gathered for the social network under surveys.

These techniques are:

1.Questionnaires

2. Experiments

3. Observations

4. Interview

5. Archival Records

When humans are our target population, questionnaires are the most commonly used method. For social networks, data collection is bit challenging task. Researchers need to design a specific question to gather the ties among respondents.Many different formats are used for collecting network ties. The one we have used is "Free Recall" method.

Girls were asked to list their three best friends and rate them according to the strength of the relationship. Girls freely recalled their friends and listed their names. We targeted their friends and collected their data. For this purpose, we delivered 600 questionnaires in the hostels of International Islamic University Hostels and got almost 498 responses.

Develop Complex Network After the retrieval of data, the next step is to develop complex network model. This is again a trivial task as it demands effort and time to develop an appropriate complex network. For the development of the network, data was entered in the "Gephi" a network visualization tool. All the girls were entered as "Nodes" and their links as "Edges". The whole complex network was directed with the weighted links according to the strength of the relationship.

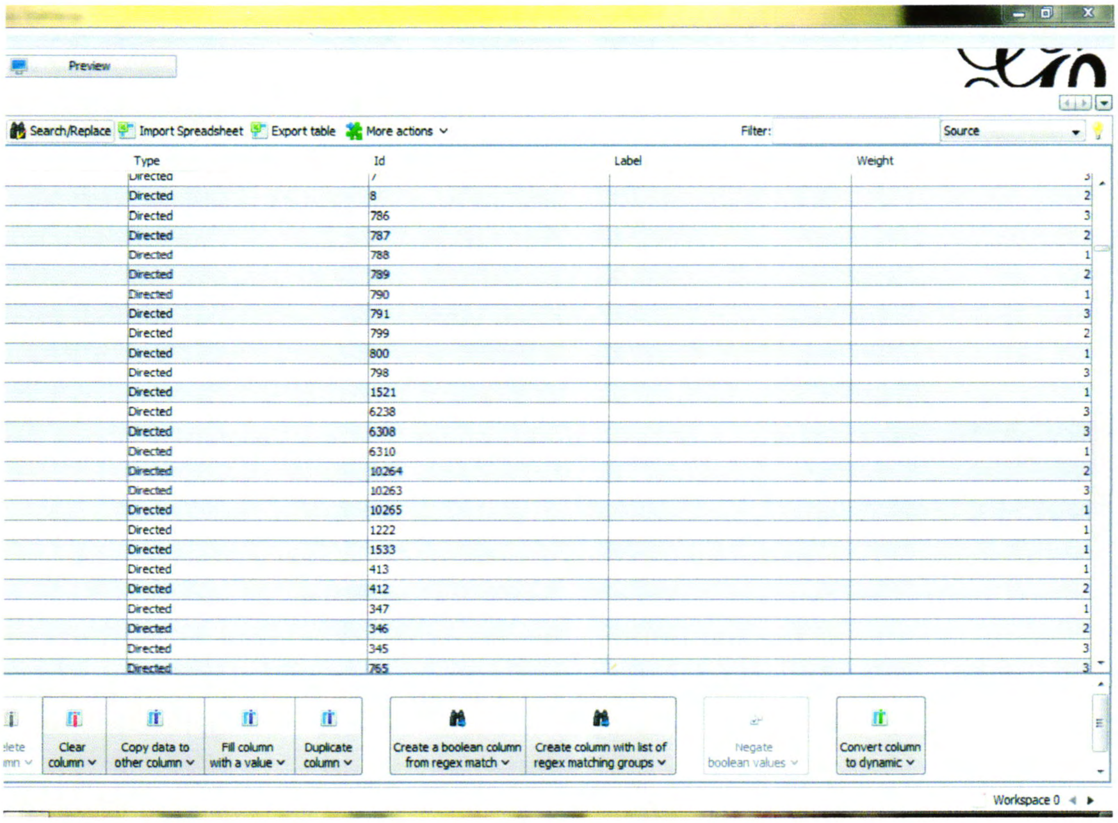


Figure 4: Gephi "A Network Visualization Tool"

Manipulate Complex Network Network Manipulation is a part of the development of a complex network. In this step, we decide suitable nodes for network construction. Network manipulation involves selection, deletion, insertion and again cleaning up the network data with unnecessary and less useful nodes and edges. After this exercise, we get the cleaned and refined network ready for the complex network analysis. The cleaned network had 498 nodes.

Perform Complex Network Analysis The next step is the analysis of the complex network. There are hundreds of possible analysis can be performed on the complex networks depending on the research goal. We have used centrality based analysis and the other ones are the statistical analysis techniques. We performed both and many others to be discussed in the later chapter.

Discover Emergent Patterns This is the last step of this framework. The basic purpose of the study of the complex system is to understand the complex nature of the different processes. This study of the complex systems and the development of the complex networks basically leads towards the perception of interesting patterns in its whole form rather than constituent.

Here is the detail of social network analysis techniques applied on the data in order to analyze behaviors .

.

Social Network Analysis As a complex network; we developed a social network of friends, attributes of the edges/ links represent their ties and the level of emotions for the friends. For example, a girl chose her three friends according to the strength of their relationship and it was represented in a network. This is done by assigning the higher value to the link, connecting two best friends. The lower and the lowest values to the second and third best friends respectively. All the relationships are represented as directed graphs. So the network has two main components "nodes" and "edges".

For analyzing network, we have conducted structural analysis.

Network Size Network size is referred as the number of nodes (i.e, people) in the network. This is the first measure to understand the structure of the network. The second measure is the number of edges connecting these nodes. Then we need to check either graph is directed or undirected. If friends are showing their choices and sentiments towards other friends then the graph would be directed. As the direction will indicate the direction of ties among friends.

Our Data Set:

Data Set	
No. of Nodes	498
No. of Edges	1226
Graph Type	Directed

Cluster Size A cluster is a group of people who are strongly connected with each other. It's been proved by many researchers that the real world networks are the strongly clustered than the other random networks. Our data set consists of 64 clusters. Average clustering coefficient indicates how nodes are embedded in their neighborhood. The average clustering coefficient of our data set is0.529 which shows that our network is highly clustered. In our network, this shows that three selected friends by a focal friend are also friends with each other. Cluster analysis highlights the people in the graph with common habits and choices.

Network Density Graph density is a measure of connectivity of a whole network. It is calculated by using equation

Graph Density = Total no. of edges/ Max. possible no. of edges

Graph Density of our data set is 0.005.

Degree The degree is the number of contacts of the node in a network. Average degree indicates the average of the degree in a complete network.

Average degree of our network is 4.9196

If the network is directed; a node might have different in-degree and out-degree measures.

In-Degree In Degree is the number of arcs/edges directed towards nodes. In our data set, a girl who is more popular in the network have the highest in-degree as the maximum number of people have chosen her as their friend.

Highest In Degree value of our Data set is 24.

Out-Degree Out Degree is the number of arcs coming out from the node. In our Data set, girls were asked to choose only three friends, so every node's out degree is 3.

Component Components of a network show the level of connectivity of a network.

Weak Component The weak component is the largest number of actors who are connected with each other irrespective of the direction of ties in the network.

Our data set contain 1 weak component.

Strong Components Strong components are the components of the network who are strongly connected with each other. For example, X is friend with Y and Z (who are also friend with each other). Meanwhile, X is also friend with A and B (who are also friend with each other). So Y and Z make the weak component with A and B as they are not directly connected to each other where as X,Y,Z and X,A,B make separate strong components. Our data set contain 210 strong components.

Diameter The diameter of the network is the observation of shortest distance between two distant nodes. This shows the distance between two farthest nodes.

Network Diameter:

Network Diameter	
Diameter	12
Average Path length	5.34
No. of shortest path	28937

Modularity Modularity measures the strength of the division of the network into modules (also known as groups, clusters or communities). Network with greater modularity has the thick connection between nodes within modules but less connections between nodes in different modules having low modularity. Overall modularity of our data set is 0.925 whereas 57 number of communities are observed.

Geodesic Distance: Geodesic distance also known as shortest path can be defined as shortest degree of separation between two nodes.

Cores Cores are the denser regions of the networks. So they help to find cohesive subgroups with K-cores partition, we can easily identify low cores of the network. This is helpful in identifying the densest parts of the network, we have 6 cores in the network.

Centrality Based Analysis To measure the location of each actor, one should understand the concept of centrality and other related concept.

Centrality measures are used to identify the actors in the network having different special characteristics. These are the most important widely used metrics to analyze the structure of the nodes in the network. Centrality refers to the position of individual nodes within network whereas, centralization is the characteristic of the whole network.

Degree Centrality The average degree of our network is 4.91 whereas degree centralization is 31.16. Input degree centralization is 0.043 and output degree centralization is 0.091.

Betweenness Centrality According to the Freeman's approach betweenness centrality is defined as the proportion of times that the node acts as a bridge between two different nodes for sending information.

Betweenness centralization of a network is 0.096.

Closeness Centrality Closeness Centrality referred as a how far a node from all others nodes in the network" [35]. If we consider the distance of node as a farness of the node. The closeness would be "the inverse of the farness".

Closeness centralization of a network is 0.187. Closeness metric ranges from 0 to 1.

Eccentricity Centrality The node having low eccentricity is considered to be the most proximal node in the network and so as easily reachable by others [65] and can be more influential than others. The detailed results are shown in the results section.

Eigenvector Centrality Eigenvector centrality focuses on the fact that a node having the connection to high centrality nodes are more important than the node having links with low centrality nodes [16]. This is a measure for finding the influential node in a network. Results are discussed in the next chapter.

After finding a unique node of highest eigenvector centrality, we applied an ego centered analysis on the ego (unique) node and her friends to analyze their behavior.

Egocentric Density of a vertex To calculate ego-centric density of ties among its neighbors, we extract the sub network of neighbors from the overall network with maximum distance 1. Two partitions were created, Ego (Focal/unique node) in class 0 represented in the figure in sky blue color, and neighbors in class 1 in yellow color 12b.

Extracted network of neighbors of unique vertex having highest eigenvector centrality and regarding this centrality measure, this node is considered to be the most influential node in the network.

After this network analysis and visualization we have performed statistical analysis on the ego centric network to analyze the influence at the first degree of separation.

4 Formal Modeling Framework

4.1 Introduction

Formal specification is a type of formal methods used for mathematical modeling of real world system [32].

For modeling a system, it is necessary to ensure that all components and their interactions are modeled correctly and at the desired level of abstraction. Formal methods are very demanding and effective strategy to use for the correct implementation of the system under study.

Z is an ISO standard formal specification language developed by programming research group in 1970's at Oxford University [77]. Formal semantics of Z specification includes mathematical notations. These notations are represented using the formal structure known as schema in Z specifications.

Z based formal specification model for modeling CAS has been used as part of FABS, an established modeling framework which is previously used by researchers [70] to effectively model and analyze the different types of CAS. FABS modeling framework involves two phases of model construction, the first level is formal model of CAS using a mathematical framework of formal specifications written in Z, and the second level is Agent-Based Model which is based on formal specifications and this is developed as a proof of concept.

We have adopted the first level of this framework to model the human behavior.

In this chapter we have presented the written specifications.

4.2 Formal Framework

We have modeled the behavior of girls participating in a network study. Formal specifications are used to specify their behavior in a clear and unambiguous way.

Sets

Sets are used in the model to declare the user defined types in the model.

The set "GIRL" is the set of all girls in the hostel.

[GIRL]

The Scenario

Girls come in the hostel, make friends and spend time with each other. Consequently either get influence of their behavior or not.

Axiomatic Definition

There is a limit of the girls who live in the hostel. We define "maximum" as maximum number of girls who can be in the hostel.

Here we have set the size arbitrarily (for no special reason) to 500 even though its actual size is not relevant.

$$
\begin{array}{|l}
maxHostelSize : \mathbb{Z} \\
\hline
maxHostelSize = 500
\end{array}
$$

System State Scenarios

We are modeling the hostel social system in these state scenarios.

- "*Joined*" is the set of girls who have joined the hostel.
- "Friends" is a subset of "joined set" and represent the set of friends who have joined and made friends.
- Those friends who have joined the hostel and take influence of the friends are represented as "influenced".
- Those who made friends but did not take influence are represented by "uninfluenced".
- The size of the girls in the hostel is maximum.

$$
\begin{array}{|l}
\hline
\textit{Hostel} \\
joined : \mathbb{F}\ GIRL \\
friend : \mathbb{F}\ GIRL \\
influenced : \mathbb{F}\ GIRL \\
\hline
\#joined \leq maxHostelSize \\
friend \subseteq joined \\
influenced \subseteq friend \\
\hline
\end{array}
$$

The Initial State

The initial state of system is declared here. In predicate part it shows that there are no girls who joined hostels, made friends and get influence. So they are represented with empty sets.

$$
\begin{array}{|l}
\hline
\textit{InitHostel} \\
Hostel \\
\hline
joined = \emptyset \\
friend = \emptyset \\
influenced = \emptyset \\
\hline
\end{array}
$$

Join the Hostel

A girl may join the hostel, if the hostel isn't already full and if the girl hasn't already joined, girls can't be friends.

ΔHostel represents the change in state of the "hostel". Set GIRL is the input. Number of joined girls are less than the maximum seats of girls in the hostel.

So all the girls have not joined the hostel. The girls who have joined the hostel have transferred to the set joined'.

Join

$\Delta Hostel$
$girl? : GIRL$

$\#joined < maxNumberOfGirlsInHostel$
$girl? \not\in joined$
$joined' = joined \cup \{girl?\}$
$joined' = joined$

Girl

Every girl is a subset of joined but all girls are not friends.

A girl is transferred to the friend set provided they spend time with other girls and made friends. Every girl must also be in joined. Joined remained unchanged

Girl

$\Delta Hostel$
$girl? : GIRL$

$girl? \subseteq joined$
$girl? \not\in friend$
$friend' = friend \cup girl?$
$joined' = joined$

Friend

Every friend is a subset of girl but every friend is not influenced. Some girls who spent time with other girls get influenced. Girl set remain unchanged.

$$
\begin{array}{|l}
\hline
\textit{Friend} \\
\Delta Girl \\
friend? : GIRL \\
\hline
friend \subseteq girl \\
friend \not\subseteq influenced \\
influenced' = influenced \cup girl? \\
girl' = girl \\
\hline
\end{array}
$$

Influenced

Every influenced is a subset of friend set but influenced is not equal to uninfluenced. Uninfluenced remain uninfluenced whether they spent time with friends or not.

Friend set remain unchanged.

$$
\begin{array}{|l}
\hline
\textit{Influenced} \\
\Delta Girl \\
influenced? : GIRL \\
\hline
influenced \subseteq friend \\
influenced \not\subseteq uninfluenced \\
uninfluenced' = uninfluenced \cup friend? \\
friend' = friend \\
\hline
\end{array}
$$

Free type Definitions

Previously we were concerned with the simple, straight forward, no problem scenario. For example, we didn't concern ourselves with the possibility that the hostel is full and so no one can join the hostel.

We ignored the possibility that a girl who doesn't join the hostel could be transferred to the influenced set.

We now address these error scenarios.

First of all we draw up a table of predictions, the set of states for which successful outcomes are defined. We add on to that table the conditions for failure.

Schema	Pre-Condition for Failure	Condition for Failure
Join	#joined <max hostel size Girl?∉joined	Hostel full: #joined ≥ max hostel size Already joined: girl?∈joined
Friend	Girl?∈joined Girl?∉friend	Not joined: girl?∉joined Already friends: girl?∈friend
Influenced	Girl?∈influenced	Not influenced: girl?∉influenced

Success and Error Schema

The success schema indicates successful outcome.

Success

Success

$report! : REPORT$

$Report! = ok$

Hostel full

This schema is defined for each error case. If the hostel is already full or mistakenly number of joined girls exceeded from the maximum hostel strength.

HostelFull

$\Xi Hostel$

$Report! : REPORT$

$\#joined \geq maxNumberOfGirlsInHostel$

$Report! = hostelfull$

Already Joined

A girl can't join again if they have already joined. If the set of all the girls who have joined, is equal to joined set. Then the one who have already joined are reported as "already joined". These could be the one who are not included in the network.

Alreadyjoined

$\Xi\ Hostel$

$girl? : GIRL$

$report! : Report$

$girl? \in joined$

$Report! = already\ joined$

Not Joined

If a girl did not join the hostel, she can't be friend with others and can't be influenced. So the girl who is not equal to the "joined" set. She is reported as "not joined".

NotJoined

Ξ *Hostel*
girl? : *GIRL*
report! : *REPORT*

---

girl? $\notin$ *joined*
Report! = *not joined*

Not Friends

If the girls, who have joined the hostel are not friends. Then they will be reported as "not friends"

NotFriends

Ξ*Hostel*
girl? : *GIRL*
report! : *REPORT*

---

girl? $\notin$ *friends*
Report! = *not friends*

Not Influenced

If they are not influenced then they are reported as "not influenced".

NotInfluenced

Ξ*Hostel*
girl? : *GIRL*
report! : *REPORT*

---

girl? $\notin$ *influenced*
Report!= *notinfluenced*

Then, in a free type definition, we define REPORT to be the set of values that describe either a schema's success or the reason for failure.

REPORT: : = Ok |Hostel Full | Already Joined | Not Joined | Not Influenced

5 Results and Discussion

In this chapter, we are going to discuss our results which we have explored in our two separate case studies.

5.1 Case Study 1

5.1.1 Statistical Summaries

Variables	Minimum	Maximum
Age	20	32
Computer Usage Hours	4 to 6hrs	1 to 3hrs
No. of Mobile Sims	1	3
Health Status	Fat	Fit
No. of Mobile Contacts	29	500
Cloth purchase/yr	15	12
Clothing Purchase Season	Winter	Summer
Online Purchase	Yes	No
Priority for Cloth Purchase	Price	Quality
Shopping Partner	Alone	Family
Matching Outfit Item	No	Rarely
Inspiration for Cloth Purchase	Celebrity	Style in Stores
1st Brand Priority	Ego	Gul Ahmed
2nd Brand Priority	Gul Ahmed	Nishat Linen
3rd Brand Priority	Al-Karam	Gul Ahmed

Table 1: Statistical Summaries of Case Study1

Discussion

As we have calculated the statistical summaries of the data, we may noted that the age in the data set ranged from 20 to 32 as plotted in the graph (5). Most people take the greatest inspiration for a new cloth purchase from the style in stores. Most people used to spend more on clothing in summer and used to go for shopping with family and prefer quality when purchasing clothes. Most people do not buy online and sometimes buy a matching outfit item with new clothes. Most people's first brand priority for buying clothes is Gul Ahmed. Second brand priority is Nishat Linen and the third brand priority is also Gul Ahmed.

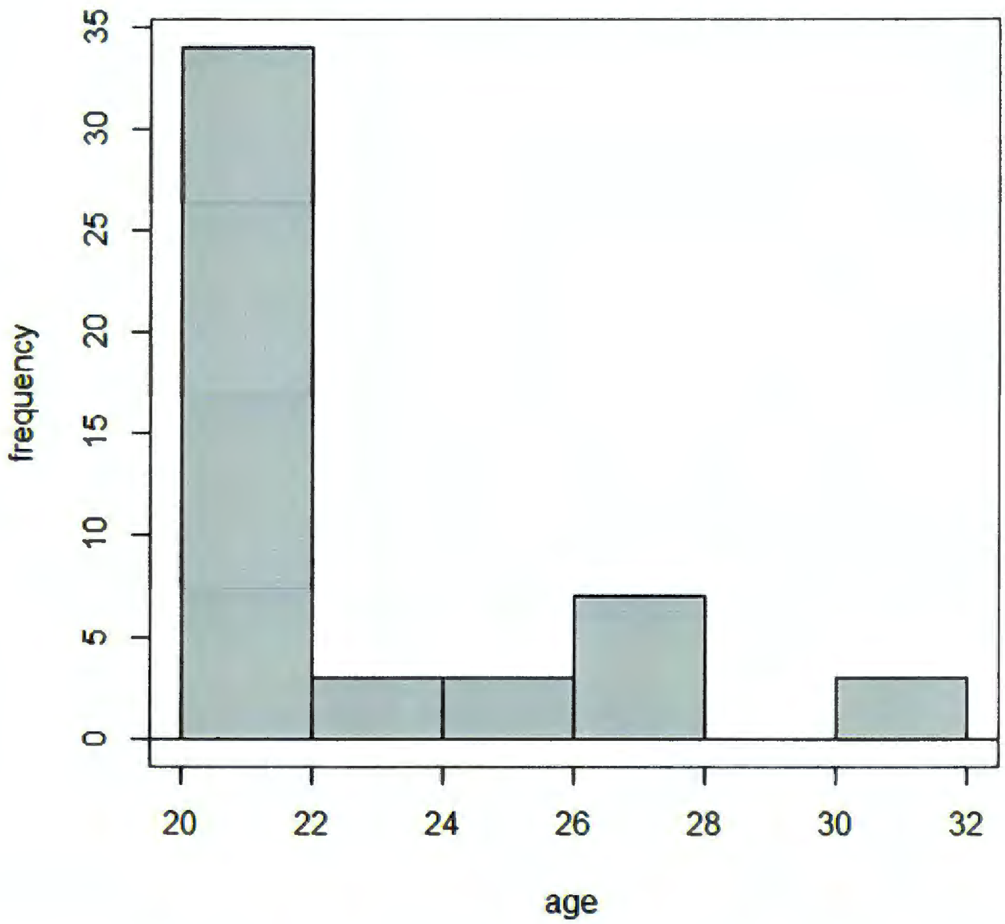


Figure 5: Histogram of Age

5.1.2 Correlation Results

Objective To find the correlation between age and interest in buying ready-made dresses.

Result

Pearson's product moment correlation.

Data: Age and Interest in buying Ready-made dresses.

$t = -0.696$, $df = 48$, $p - value = 0.24$

Alternative hypothesis: True correlation is less than 0.

95% CI : $-1.0, 0.13$

Sample Estimate : Cor -0.09995428

Discussion Since the correlation shows the interdependence between two independent variables. The above-mentioned result shows ($r = -0,099$) that there is weak negative relationship among the above mentioned variables. 95% CI shows that we are 95% confident that population parameter (rho) lies between $(-1, 0.3)$

Objective To find correlation test between Age and Interest in buying cultural dresses.

Result

Pearson's product moment correlation.

Data: Age and Interest in buying cultural dresses.

$t = -1.5$, $df = 47$, $p - value = 0.068$

Alternative hypothesis: True correlation is less than 0

95 CI : -1 , 0.02

Sample Estimates : Cor -0.215 (Weak negative relationship)

Discussion : Since the correlation shows the interdependence between two independent variables, so the above-mentioned result shows ($r = -0.215$) that there is weak negative relationship among the above mentioned variables. 95% CI shows that we are 95% confident that population parameter (rho) lies between $(-1, 0.02)$.

Objective To find the correlation between Age and Interest in buying unstitched clothes.

Result

Pearson's Product moment correlation

Data: Age and Interest in buying unstitched clothes.

$t = -0.40$, $df = 48$, $p - value = 0.34$

Alternative hypothesis: True correlation is less than 0.

95% CI= -1.0 ,0.17

Sample Estimates: Cor -0.058 (Weak negative relationship)

Discussion Since the correlation shows the interdependence between two independent variables, so the above-mentioned result shows ($r = -0.058$) that there is weak negative relationship among the above-mentioned variables. 95% CI shows that we are 95% confident that population parameter (rho) lies between $(-1, 0.17)$

Objective To find the correlation between the Age and Interest in buying western dresses.

Result

Pearson's Product moment correlation

Data: Age and Interest in buying western dresses.

$t = -0.6005$, $df = 48$, $p - value = 0.2755$

Alternative hypothesis: True correlation is less than 0.

95% CI = -1.00, 0.152

Sample Estimates : Cor -0.086 (weak negative relationship)

Discussion Since the correlation shows the interdependence between two independent variables, so the above-mentioned result shows ($r = -0.086$) that there is weak negative relationship among the above-mentioned variables. 95% CI shows that we are 95% confident that population parameter (rho) lies between $(-1, 0.152)$

Objective To find the correlation between Age and No. of contacts in your mobile.

Result

Pearson's product moment correlation

Data: Age and No. of contacts in your mobile.

$t = -o.8152$, $df = 48$, $p - value = 0.2095$

Alternative Hypothesis: True correlation is less than 0.

95% CI : -1.00, 0.121

Sample Estimates: Cor -0.1168 (weak negative)

Discussion Since the correlation shows the interdependence between two independent variables, so the above-mentioned result shows ($r = -0.1168$) that there is weak negative relationship among the above-mentioned variables. 95% CI shows that we are 95% confident that population parameter (rho) lies between$(-1, 0.121)$

Objective To find the correlation between Age and No. of mobile sims.

Result

Age and No. of mobile sims.

$t = 0.8219$, $df = 48$, $p - value = 0.79$

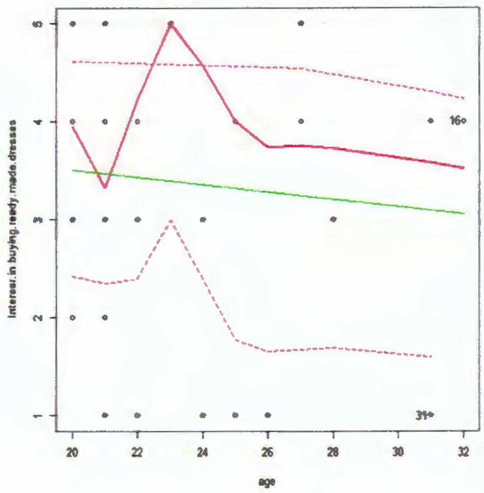
Alternative Hypothesis = True correlation is less than 0.

95% CI $= -1.00, 0.343$

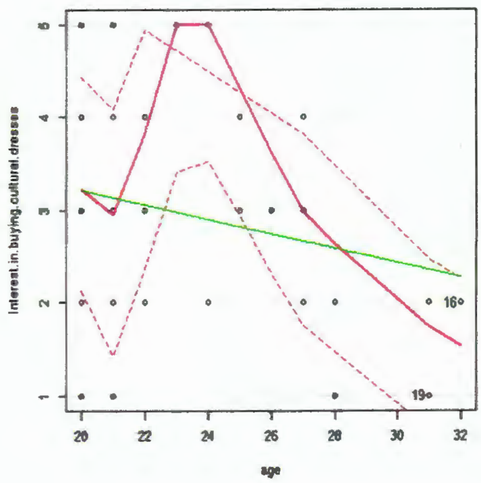
Sample Estimates: Cor 0.1178 (Weak positive)

Discussion Since the correlation shows the interdependence between two independent variables, so the above-mentioned result shows ($r = 0.1178$) that there is weak negative relationship among the above-mentioned variables. 95% CI shows that we are 95% confident that population parameter (rho) lies between $(-1, 0.343)$

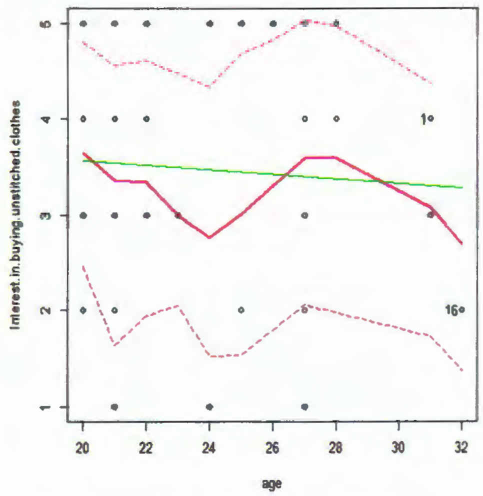
In examining the pattern of correlation between age and different categories of dresses we find that;



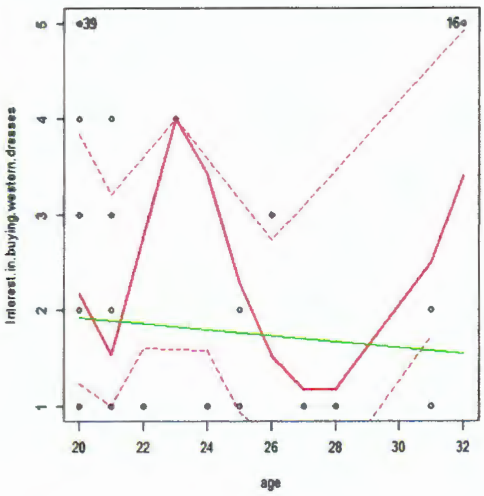
(a) Interest in Ready made Dresses



(b) Interest in Cultural Dresses



(c) Interest in Unstitched Dresses



(d) Interest in Western Dresses

Figure 6: Correlation Results

5.1.3 Regression Results

Objective To find dependency of 1 variable upon other variable.

Method
Click on "Statistics" in Rcmdr → "fit model" → "linear regression".
Results
1m (Formula = Interest in buying ready made dresses ~ Age, Data = Name of data file)
Residuals:
Min = −2.4640
IQ = −0.4640
Median = 0.4988
3Qu = 0.7593
Max = 1.7593
Coefficients:

	Estimates	Std Error	t.value	pr(>|t|)
Intercept	4.245	1.227	3.45	0.00115
age	-0.03722	0.05348	-0.696	0.48979

Residual standard error : 1.252 on 48 degrees of freedom
Multiple R-squared : 0.009991
Adjusted R-squared : −0.01063
F-statistics : 0.4844 on 1 and 48 DF.
P-value= 0.4898

Discussion Linear regression model shows interest in buying ready made dresses regress negatively with age with the value of slop(−0.03722) for a unit change in age, the dependent variable (i.e, interest in buying ready made dresses decreases with the rate of 3.722). Residual in results showed the difference between the expected value and the estimated value. The value of multiple R-squared shows the variation in the dependent variable explained by regression coefficients(values of slope).

Objective To find dependency of 1 variable upon another.
Results
1m (Formula = Interest in buying cultural dresses~age, Data= Name of Data File)

Residuals
Min = −2.22
IQ = −1.06
Median = −0.14
3Q = 0.85
Max = 2.09
Coefficients:

	Estimate	Std.error	T.value	Pr(>|t|)
Intercept	4.81183	1.21316	3.966	0.000248
Age	-0.07955	0.05271	-1.502	0.137922

Residual standard error: 1.225 m 47 degree of freedom (one observation is deleted due to missingness)
Multiple R-squared : 0.04623,
Adjusted R-squared : 0.02593.
F-statistic : 2.278 on 1 and 47 DF.
P-value : 0.1379

Discussion Linear regression model shows interest in buying cultural dresses regress negatively with age with the value of slop (−0.07955) for a unit change in age, the dependent variable (i.e, interest in buying cultural dresses decreases with the rate of 7.955). Residual in results showed the difference between the expected value and the estimated value. The value of multiple R-squared shows the variation in the dependent variable explained by regression coefficients (values of slope).

Objective To find dependency of 1 variable over another.
Results
1m (Formula = Interest in buying unstitched clothes~age, Data= Name of Data file)
Residuals:
Min = −2.5408
IQ = −1.1010
Median = 0.4473

3Q = 1.435

Max = 1.6252

Coefficients:

	Estimate	Std.error	T.value	Pr(>|t|)
Intercept	4.03890	1.3443	3.004	0.00422
Age	-0.002372	0.05855	-0.405	0.68721

Residual Standard Error : 1.371 on 48 degree of freedom

Multiple R-squared : 0.003407,

Adjusted R-squared : -0.0173

F-statistics : 0.1641 on 1 and DF 48.

P-value : 0.6872.

Discussion Linear regression model shows interest in buying unstitched regress negatively with age with the value of slop (-0.002372) for a unit change in age, the dependent variable (i.e, interest in buying unstitched clothes decreases with the rate of 0.2372). Residual in results showed the difference between the expected value and the estimated value. The value of multiple R-squared shows the variation in the dependent variable explained by regression coefficients (values of slope).

Objective To find dependency of 1 variable over another.

Results

1m (Formula = Interest in buying western dresses~age, Data = Name of Data file)

Residuals:

Min = -0.925

IQ = -0.8941

Median =-0.7054

3 Qu =0.9109

Max = 3.4519

Coefficients:

	Estimate	Std.Error	T.value	Pr(>|t|)
Intercept	2.55466	1.20262	2.124	0.0388
Age	-0.03146	0.05238	-0.601	0.5510

Residual standard error : 1.226 on 48 degrees of freedom

Multiple R-squared :0.007457.

Adjusted R-squared : -0.01322

F-statistics : 0.3606 on 1 and 48 DF.

P-value : 0.551

Discussion Linear regression model shows interest in buying western dresses regress negatively with age with the value of slop (-0.03146) for a unit change in age, the dependent variable (i.e, interest in buying western dresses decreases with the rate of 3.146). Residual in results shows the difference between the expected value and the estimated value. The value of multiple R-squared shows the variation in the dependent variable explained by regression coefficients (values of slope).

Comprehensive Discussion of Linear Regression We conducted a linear regression to model the relationship between age (x) and several dependent variables. The results of linear regression analysis support the correlation results, these results showed that interest in buying ready-made dresses, cultural dresses, unstitched clothes, western dresses all regress negatively with age with the percentage of 3.7, 7.95, 0.23, 3.14, 44 respectively

5.1.4 Interpretations of Correlation and Regression Results

The Figure 6a shows people of age 21 showing least interest in ready-made dresses where people of age 23 is showing too much interest in ready-made dresses.

After the age of 23, people's interest in ready-made dresses is dropping significantly till the age of 26 and then from 26 to 32 their involvement in ready-made dresses almost remains the same. So the people of age 21 and 23 showing the significant difference in their choices for ready-made dresses.

Students of age 21 are normally considered as newcomers in the undergraduate programs. Students living in the hostels usually come from the rural areas of Pakistan having less interest in clothes and fashions due to lessening media and social exposure. But with every passing year their exposure to urban areas make them more involved in such activities like clothing interest and technology awareness. After ending their graduation studies in the average age of25, students of age 26 or more are enrolled in MS and Ph.D programs. So according to the analysis results, it is observed that people enrolled in the higher educational programs are losing their interests in such activities.

The figure 6b shows at the age of 21, again people are showing very little interest in cultural dresses whereas students from the age of 23 to 25, they are showing most interest in cultural dresses and again after 25, there is a significant drop of the line up to the age of 32.

According to the figure 6c there is a weak negative relationship between age and their interest in buying unstitched clothes is observed, even people of age 23 and 24 are showing least interest and people of27 and 28 years of age showing bit more interest comparatively.

On average people of all age groups are almost equally involved in unstitched clothes as no significant variation in their choices is observed. In figure 6d the curve shows people of age 21 are showing least interest in western clothes. At the age of 23 they are again showing much interest in western clothes. After the age of 23, there is again a significant drop of the curve till the age of 27. Then again after 27, people's interest in western dresses increased up to the age of 32 but not more than the people of age 23.

Analysis of the collected data revealed that students of age 23 living in the hostel are so involved in dresses. So the significant finding of our research is that younger students in the hostels are not much interested in fashion activities but as they belong to the age group having a lot of energies, motivation, and adaptability to attract and learn new things they get involved in such activities in a year or so but after 25 years of age, students involved in higher education programs start losing their interest in these activities again.

Conclusion By our conducted case study results, we have concluded that this decline in the dressing interest could also be a result of circumstances students face due to their tough studies. Secondly this study also validates the idea of previous research , as with an increase in age people's interest in different types of dressing starts losing. We have also observed their choices and interests variations in different age groups but as an emergent behavioral pattern we have discovered the loss in interest. This study is considered as an example of the complex system which we have dug to find the emergent patterns. We have discovered interesting behavioral patterns of decline in dressing interest with the growing age.

5.1.5 Previous Studies Correlated with our Findings

In the current study (Case study 1), we examined the correlation of dressing interest and analyzed the effect of age on these interests. In the previous studies where a lot of research has been done on age effects over different life interests; we reviewed

Age and Clothing Interests As similar to our studies Arpita, Ankita, ceeba and Rajlaxmi (2011) [42] made a hypothesis in their research that age doesn't affect on fashion clothing in Indian society. The hypothesis was rejected and their results showed that age is a major determinant of their involvement in fashion clothing and young people are more involved in fashion wear than old people. They noted the fact, that this may be due to the increased awareness and availability of international brands in India.

In the same way, our research is also supported by the earlier research which shows that purchaser's fashion clothing interests are significantly affected by age [29] [55] [56].

In another study Kozar (2012) [44] noted that female clothing purchase behavior and other products are affected by the model's age because the people highly rated the older models, also their purchasing habit for clothing depend upon their similar body size and looks with the model.

In another research by Katz and Lazarsfeld (1955) [41] and Tyrchniewiez and Gonzales (1978) [72] showed that female interest in clothing decreased with growing age.

On the contrary Lee et al.(1997) [45] took a data set of elder people more than 60 years of age and reported that elder women who are alone spend more on clothing than couples and women are always more attracted to spend on apparel than men. They also considered the fact that it may be due to the fact that elder people after retirement get more time for apparel expenditure.

Ebeling and Rosencranz (1961) [23] and Shipley and Rosencranz (1962) [69] also reported elder women clothing interests.

Many researchers reported some other reasons of clothing interests, Hertman (1949) [36] noted in his research that as personality develops the interest in clothing diminishes so the clothing importance varies with levels of personality development. His research was supported by Aiken (1963) [1] who noted that non self-actualizing people have the higher interest in clothing. Ryan (1953) [64] found in his research that people who have the higher interest in clothing are more dependent on environment for their adjustment whereas, people who are less interested in clothes are self-directed people.

On the contrary, some researchers showed the positive aspects of clothing interests. Like Sharma (1980) [68] noted that clothing interest is a reason of social orientation. Rosencraz (1972) [61] in his research stated that females, students of social sciences, home economics, arts and humanities have the higher interest in clothing. Rosencranz (1960) [60]reported that people of urban areas and of higher income brackets are mostly interested in clothes.

Age and Toys Interest in children Servin et al (1999) [66] found that as the kids age increases their interest in feminine toys decreases. Sex differences were clear in choosing masculine toys and feminine toys. Boys always like masculine toys and girls like feminine toys but the girls interests in feminine toys decreases with age than boys interest in masculine toys whereas (Black more et al. , 1979; O Brien & Huston, 1985; Robinson and Morris, 1986 as cited in Servin , 1990) [11] [54] [58] showed that girls get more gender-stereotyped with age, that's the reason they get involved in girlish things with growing age.

Age and Computer Interest Ellis and Allaire (1999) [24] noted in their research about age and computer interest that computer interest negatively correlates with age whereas computer anxiety is positively correlated with age.

Age and Food Preferences Cooke and Wrdle (2005) [19] in their study about age and the food preferences noted that boys in their early ages are more picky for food than girls who are more so during their adolescence and they gave the explanation for these findings are maybe girls become more picky in this age due to their diet and weight issues.

Age-Related Decline in Motivational Tasks As we reviewed over the clothing interest and age, Dickerson (1993) and fisher (1997) [21] [22] noted in their research that as the age increases even the motivational tasks become less interesting for people due to age-related decline but still older and younger people show better performance on task which are of interest than less interesting tasks.

Here we had summarized the correlation of age with different interests in the table below.

Increasing Age Effects	Clothes	Computer Interest	Computer Anxiety	Picky for Food	Motivational Task	Kids Interest in Toys
	Negative Correlation	Negative Correlation	Positive Correlation		Negative Correlation	
Young Girls				Positive Correlation		
Young Boys				Negative Correlation		
Baby Girls						Negative Correlation
Baby Boys						Positive Correlation

Table 2: Correlation of Age with different Interests

5.2 Case Study 2

5.2.1 Basic Statistical Summaries of the Data Collected

This table shows the choices of girls/ students who are living in the girls' hostel. The collected data is about their every day habits, dressing interests and food choices.

The presented table, gives us the complete picture of the habits of girls participating in the network study. The age of the girls living in the hostel ranges from 22 to 35, maximum people are 22 years old. Most people use computer 2 hours in a day and sleep 8 hours in a day.

81% people don't purchase online and very few people likes online shopping. Most people prefer "comfort" while purchasing clothes. Almost 50% people prefer dark color and 50% people like light color clothes. Mostly people like to do shopping with family. Almost 50% people buy matching outfit item with clothes and 50% don't do the same. Maximum people get inspiration from style in stores for buying new clothes and buy five times in a year. A high percentage of people purchase clothes in summer.

Attributes	Maximum People	Minimum People
Age	22 (20.32%)	35 (0.2%)
Meet. with best friend/day	Thrice a day	Twice a day
Meeting with friend/day	Thrice a day	Once a day
Meet. with acquaintance/day	Thrice a day	Twice a day
Daily computer usage hours	2hrs in a day (18.51%)	16hrs in a day (0.2%)
No. of sims	1 Sim	6 Sims
Online purchase	No (81.89%)	Yes (18.11%)
Priority for cloth purchase	Comfort (17.91%)	Material (10.66%)
Preference for dark colors	Yes (66.27%)	No (33.8%)
Preference for light colors	Yes (67%)	No (33%)
Who do you mainly shop with	Family (58.95%)	Alone (5.84%)
Buy matching outfit item	No (50.5%)	Yes (49.5%)
Int.in western dresses (1-10)	1 (42.05%)	9 (2.62%)
Int. in cultural dresses (1-10)	2 (14.29%)	1 (6.04%)
Int. in readymade dresses (1-10)	10 (15.09%)	1 (5.84%)
Int.in unstiched clothes (1-10)	8 (15.09%)	4 (5.43%)
Inspiration for new clothes	Style in stores (28.96%)	Magazines (11.27%)
Purchase of a new cloth/ year	5 times a year (36.84%)	15 times a year (12.27%)
Season of cloth purchase	Summer (44.87%)	Autumn (10.26%)
Health status	Average (43.46%)	Fat (11.47%)
How active are (1-10)	8 (20.32%)	2 (1.81%)
Freq. of exercise in a week	Don't do (55.73%)	6 days/week (1.61%)
Like fast food (1-10)	5 (17.3%)	3 (8.85%)
Like home cooked(1-10)	10 (38.83%)	9 (0.2%)
Sleep hours	8hrs/day (29.78%)	4 (0.8%)
Freq. of fast food in a week	Once a week (24.14%)	Daily (2.21%)
Freq. of fuzzy drinks in a week	Once in a week (30.99%)	Daily (3.62%)
Freq. of packed juices in a week	5 times/week (17.3%)	9times (4.43%)
Do you smoke	No (98%)	Yes (1.21%)
Freq. of smoking	Don't smoke (97.99%)	everyday (0.2%)
Freq. of eating vegetables	More than once a week (48.49%)	More than once daily (1.61%)
Freq. of eating fresh fruits	More than once a week (33.6%)	Never (2.4%)
Freq. of eating choclates	More than once a week(28.37%)	More than once daily(4.63%)

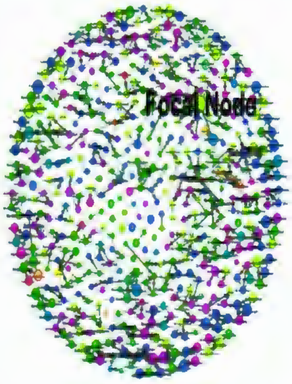
Table 3: Statistical Summaries of Case Study 2

5.2.2 Frequency Distribution of Brands Choices

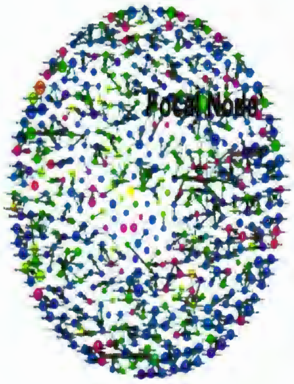
Favorite Brands Name	Frequency of People
Adidas (German)	2
Al-Karam	5
Asim Jofa	5
B.G (U.K)	1
Bareeze'	**11**
Break Out	2
Charizma	3
Chen one	2
Cynosure	2
Ego	**18**
Generations	12
Gul Ahmed	**41**
HSY	3
J.	**22**
Kayseria	6
Khaadi	**47**
Lime lite	3
Loft (U.S)	1
Mango (Spanish)	**7**
Maria B	1
Mausemmary	1
Max Mare (Italian)	1
Nine West (U.S)	1
Nishat	**33**
Origin	3
Outfitter	**9**
Polka Dot	2
Rang Ja	4
Sana Safinaz	**9**
Shahposh	6
Warda	**13**
Zahra Ahmad	3

Table 4: Frequency Distribution of most liked Brands

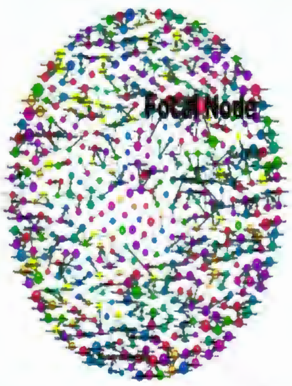
5.2.3 Visual Analysis of Girls habits living in the Hostel



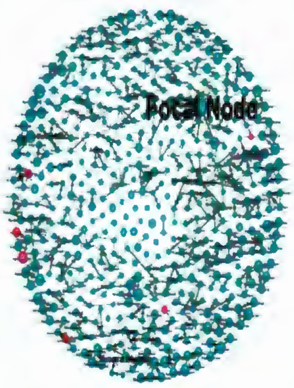
(a) Age of Girls living in the Hostel



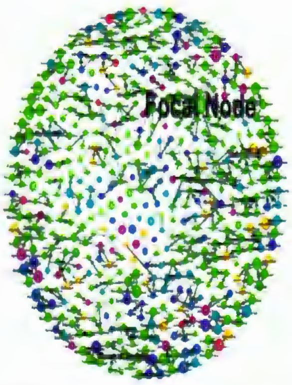
(b) Sleep hours



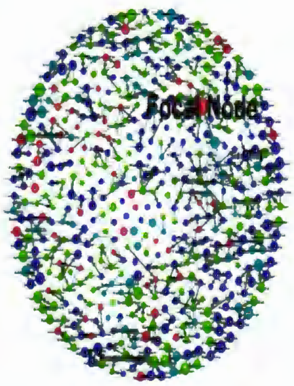
(c) Daily computer usage hours



(d) Smoking Habit



(e) Physical activity in a week



(f) Health Status

Figure 7: Visual Analysis of Attribute Data

Each circle (node) in the network model represents one person. Their size is based on their Eigenvector centrality measure. Lines (edges) between them represent their relational ties and their thickness is the representation of the strength of the relationship, thicker the edge is, stronger the relationship.

In figure 7a, the nodes in green color represent people with 22 years of age. Maximum people in the network belong from this age group.

In figure 7b nodes in blue color represent the people who sleep 8 hours in a day.

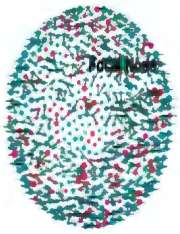
In figure 7c Nodes in red color represent the people who use computer 2 hours in a day.

In figure 7d nodes in blue color represent the people who don't smoke whereas nodes in red color are those who smoke.

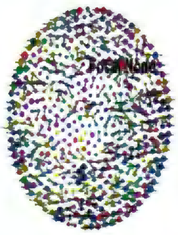
In figure 7e nodes in green color represent the people who don't do physical activity.

In figure 7f nodes in purple color are the representation of people who tell their health status as average whereas red color is the representation of obese people, green for fit and blue for weak people.

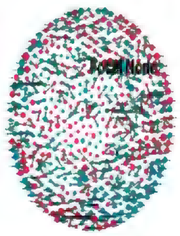
5.2.4 Visual Analysis about Dressing Choices



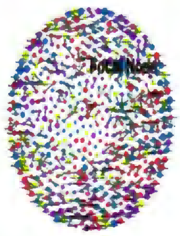
(a) People who like to wear brands



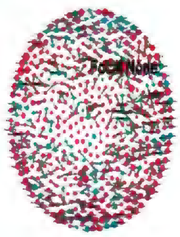
(b) Inspiration for buying new clothes



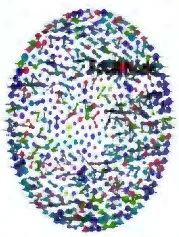
(c) Buying matching outfit items with clothes



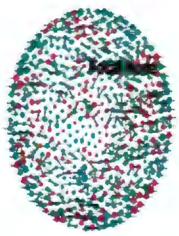
(d) New cloth purchase/year



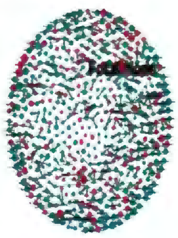
(e) Online Purchase



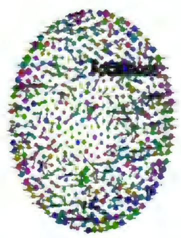
(f) Period of year spend on clothing



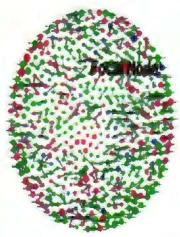
(g) Preference for dark colored clothes



(h) Preference for light colored clothes



(i) Priority when purchasing clothes



(j) Shopping Partner

Figure 8: Visual Analysis of Dressing choices

In figure 8a Red colored nodes are the representation of people who don't wear brands whereas blue colored nodes are the people who like to wear brands. In this figure 8b, purple color nodes are the people who take maximum inspiration for clothes purchase from "styles in store" . Maximum percentage of 28.97 people have given this answer. Navy blue colored nodes represent people who take inspiration from the family which are 20.12%. Mustard colored nodes take inspiration from friends, blue colored nodes are people who take inspiration from celebrities and maroon colored node takes inspiration from magazines. In figure 8c blue colored node represent people (which are 50.5%) who don't buy matching outfit item with clothes whereas red colored nodes are people (49.5%) who like to buy matching outfit items. In figure 8d Red colored node represents people who purchase new clothes 5 times in a year, these are 36% people. Blue colored nodes represent people who purchase 12 times, purple colored buy 20 times, green colored buy 15 times with the percentage of 27, 23 and 12 respectively.

In figure 8e red colored node represents people who don't purchase online and blue colored nodes represent people who like online shopping with the percentage of 81.89 and 18.11 respectively. In figure 8f purple colored nodes represent people who mostly spend on clothing in summer, blue colored in spring, red colored in winter and green colored in autumn with the percentage of people 44.87, 27.75, 17.3 and 10.26 respectively. In figure 8g blue colored nodes represent people who prefer dark colored clothes and red colored are those who do not with the percentage of 66.27 and 33.8 respectively. In figure 8h Blue colored nodes represent people who like light colored nodes whereas red colored are those who don't with the percentage of 67 and 33 respectively. In figure 8i purple colored nodes are the representation of people who prefer comfort while buying clothes, blue colored prefer quality, pink colored prefer color, maroon prefer brand, green colored prefer style mustard colored prefer price whereas parrot green prefer material with the percentage of 17.91, 16.7, 15.49, 14.89, 12.07, 12.07 and 10.66 respectively. In figure 8j green colored nodes represent people who used to go for shopping with their family, red colored with their friends and blue colored alone with the percentage of 58.95, 35.21 and 5.840 respectively.

5.2.5 Network Report

Network Report	
Context	
Nodes	498
Edges	1226
Clusters	64
Network Overview	
Strong Components	210
Weak Components	1
Average Degree	2.4
Average weighted Degree	4.79
Network Diameter	12
Average path length	5.34
No. of shortest paths	28937
Graph Density	0.005
Modularity	0.929
No. of Communities	57
Average Neighborhood Overlap	0.298
Average Embeddedness	1.488
Cores	6
Cliques	108
Node Overview	
Average Clustering Coefficient	0.529

Table 5: Network Report

5.2.6 Centrality Based Analysis

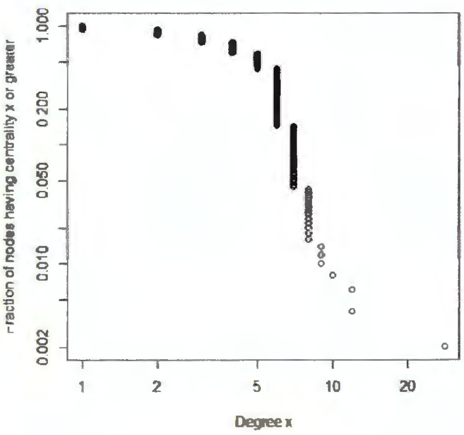
Centrality Based Analysis	
Centrality types	**Measures**
Degree Centralization	31.16
Input Degree Centralization	**0.04342**
Output Degree Centralization	0.0918
All closeness Centralization	**0.187**
Betweenness Centralization	0.096

Table 6: Centrality Measures

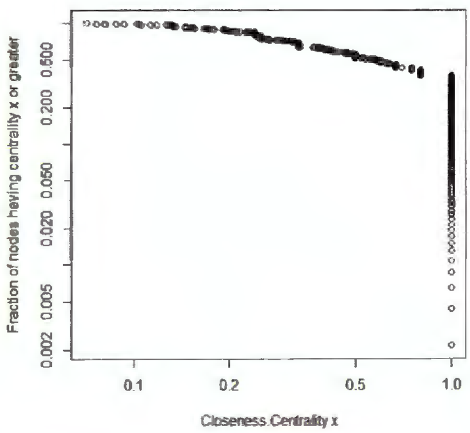
5.2.7 Powerlaw Distribution of Centrality Measures

These plots represents "the correlation of centrality metrics and the number of nodes in the network of 498 nodes.

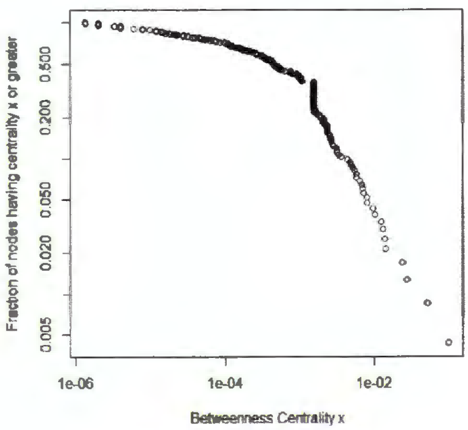
Figure 9a represents the degree of the nodes, there is only one node with degree of 24 whereas minimum degree is 1. Figure 9b shows the closeness centrality, it says many nodes have 1 closeness centrality whereas it ranges between zero and 1. Figure 9c shows the betweenness centrality of the nodes. Figure 9d represents eccentricity measure in the network, maximum observed eccentricity in the network is 18. Figure 9e showing eigenvector centrality of the node, maximum eigenvector centrality observed in the network is 1.



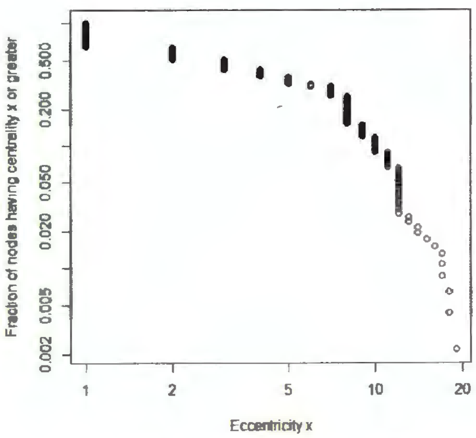
(a) Degree Centrality



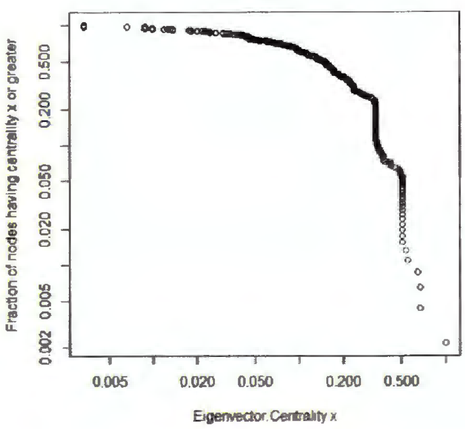
(b) Closeness Centrality



(c) Betweenness Centrality



(d) Eccentricity Centrality



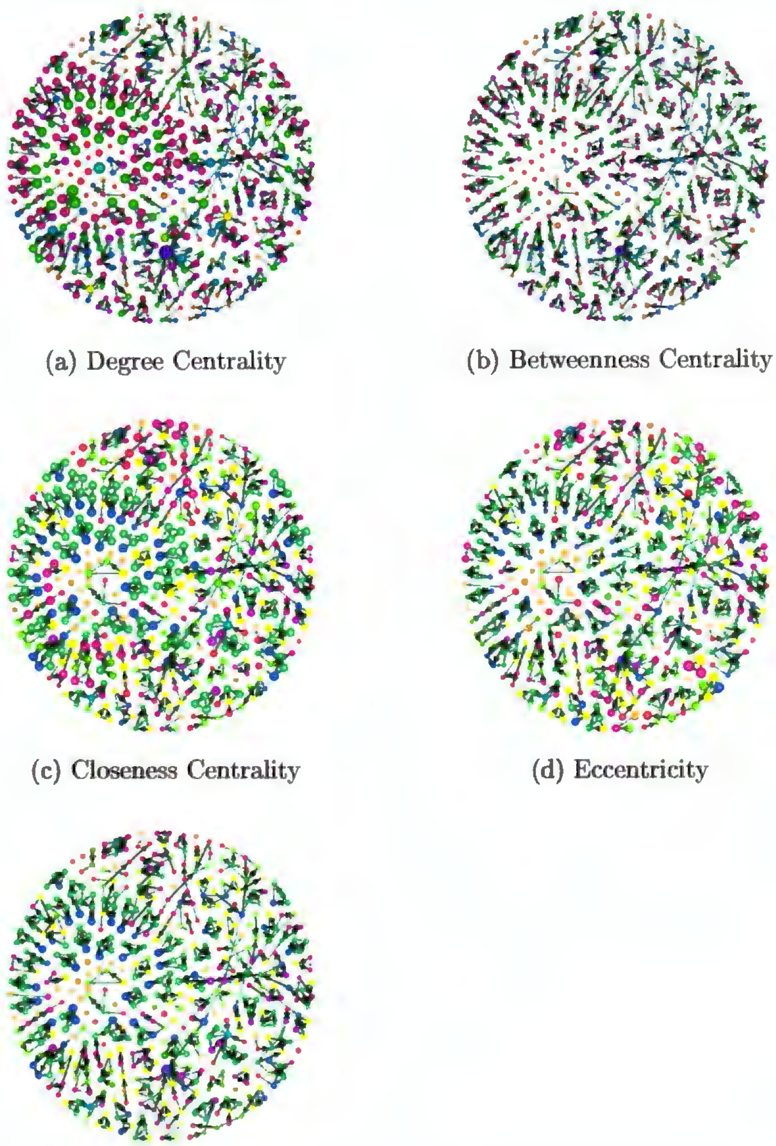
(e) Eigenvector Centrality

Figure 9: Power law Distribution of Centrality Measures of the whole network

5.2.8 Centrality-Based Visual Analysis

For Centrality-based visual analysis, we used Gephi software with "Frutcherman reingold layout". Focal node is in purple color in all networks represented in figure 12, size correlates with centrality measures whereas colors represent the degrees of the nodes in each network.

In Fig 10a, the one having purple color with the biggest size are the focal node having the highest degree in a network. In Fig 10b shows the visual display of the ranking (size) of the nodes with highest betweenness centrality and partitioned (colored) on the basis of degree. In Fig 10c shows the ranking (size) of nodes with the high closeness centrality. In Fig 10d shows the ranking (size) of nodes with high eccentricity value. In Fig 10e shows the ranking (size) of the nodes with high eigenvector centrality which is the indication of the most influential node.



(a) Degree Centrality

(b) Betweenness Centrality

(c) Closeness Centrality

(d) Eccentricity

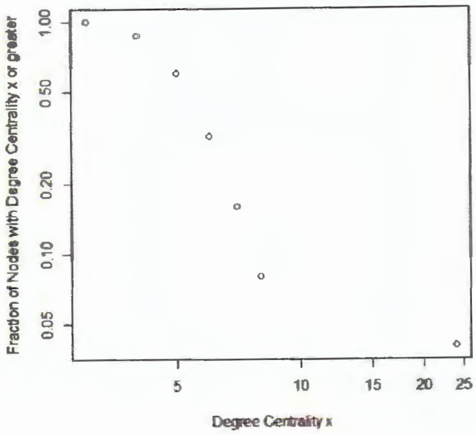
(e) Eigenvector Centrality

Figure 10: Centrality Based Visual Analysis.

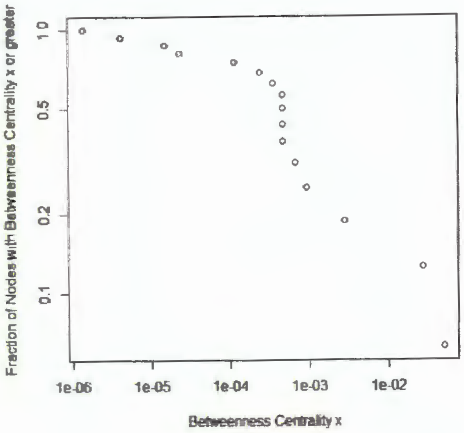
5.2.9 Powerlaw Distribution of Centrality Measures of the focal node and the neighboring friends

We have scaled these centralities and plotted them by fitting powerlaw distribution.

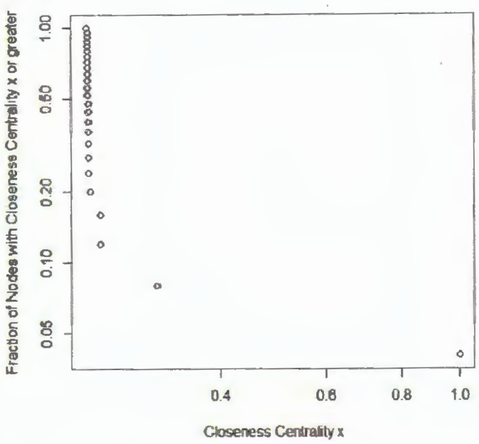
In 11a, Our focal node has a maximum degree centrality 24, all others friends' degree range from 1 to 9. In 11b, our focal node betweenness centrality is 0.05. Our focal node's closeness centrality is 0.31 represented in 11c. Our focal node's eccentricity is 7 in 11d. 11e is the representation of our focal node's eigenvector centrality which is 1 and this is the highest eigenvector centrality from the whole network.



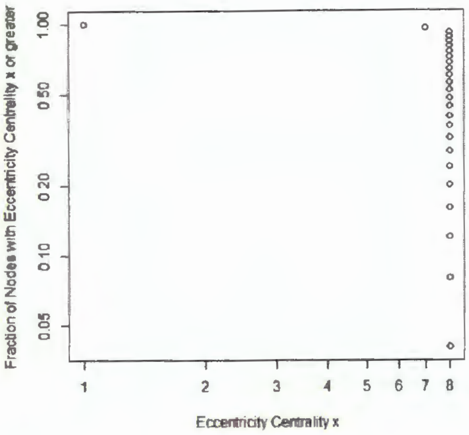
(a) Degree Centrality



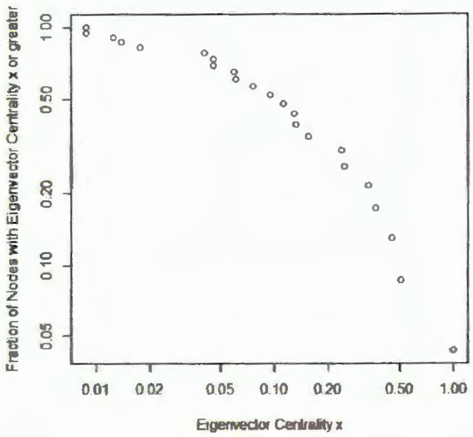
(b) Betweenness Centrality



(c) Closeness Centrality



(d) Eccentricity



(e) Eigenvector Centrality

Figure 11: Powerlaw Distribution of Focal node and the Neighbors

After having the complete report of the network, we started focusing the node having the highest degree and eigenvector centrality which proved this node as the most influential and popular node in the network. We looked for the neighbor of the node with the one degree of separation. Figure 12a and 12b shows the focal person with her neighbors in the network.

We used Kamada-Kwai algorithm to get the clear image of the neighboring nodes of a focal person.

5.2.10 Egocentric Density of a Node

To find the egocentric density of ties among node's (focal node having highest eigenvector centrality) neighbors, we have extracted the sub network from the overall network. Density calculated was 0.097 and average degree of this sub network is 5.07.

(a) Focal Node in a complete network (b) Focal Node and the Neighbors

Figure 12: Focal Node and the Neighbors

5.2.11 Behavioral Analysis Results

We performed statistical analysis over the attribute data of the focal node and her neighboring nodes (friends). We took contact type as a function of different variables such as their sleeping habit, frequency of physical activity in a week, fast food intake and fresh fruits and juices intake to analyze whether the focal node and her contact type (Best friend, friend, acquaintance) have same habits or not. Figure 13a 13b 13c 13d show the strip chart between the focal person, her contacts, and their habits. Focal node sleeps 7 hours in a day, physical activity of focal node in a week is 0. Focal node's fast food intake is 8 times in a week and fresh fruits and juices intake is 3 times in a week. These results did not show any similarity in the habits of the focal node and her first-degree friends, as well as no behavioral influence, is observed.

These results show that students living in the hostel don't take influence of each other even when they spend their whole day with each other, It could be the factor of their different backgrounds and strong family system in Pakistan. Additionally this fact is also inevitable that they might don't spend years with each other, so they can take influence.

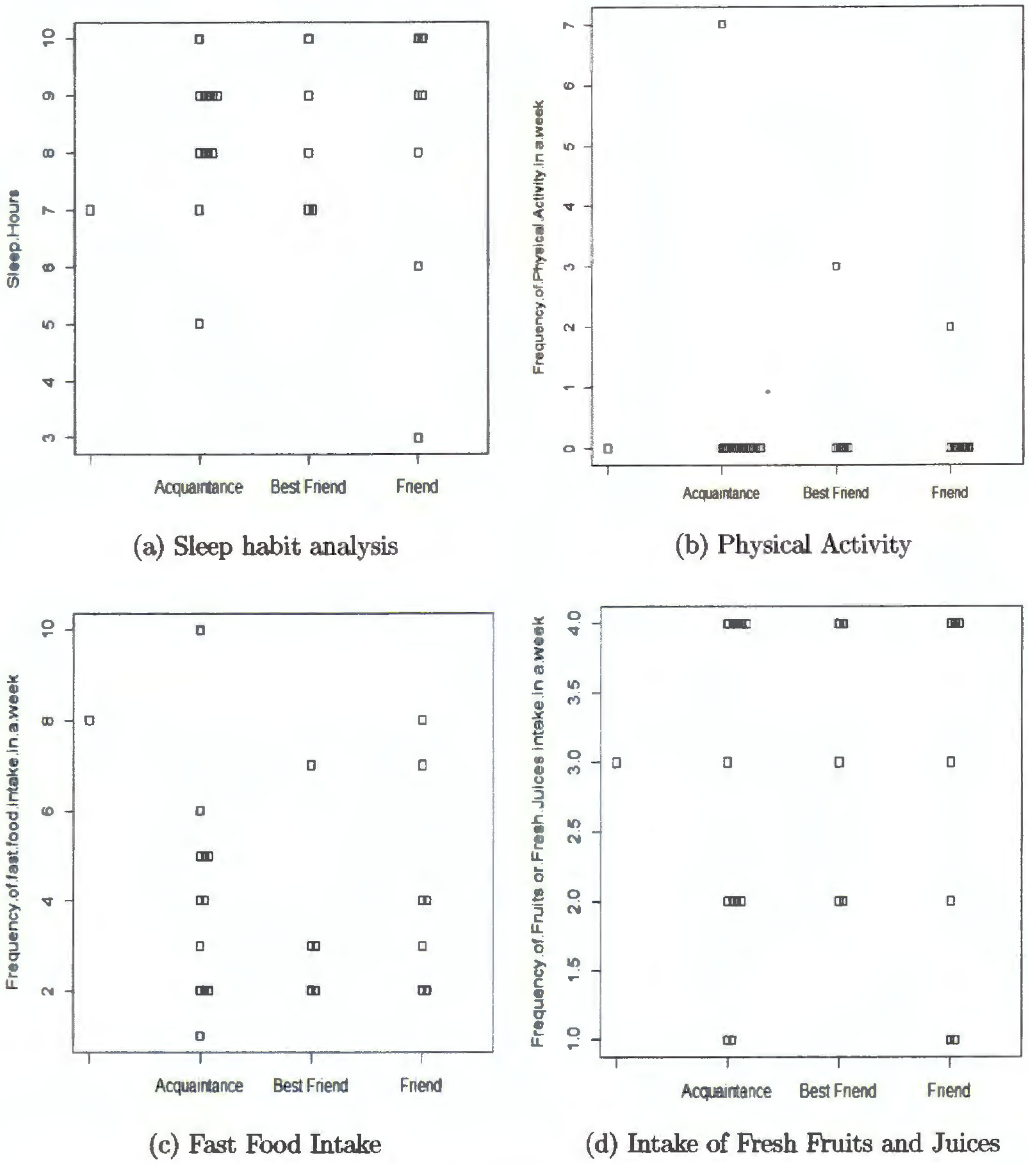


Figure 13: Behavioral Analysis using Strip Chart

5.2.12 Interpretation of the results of Case Study 2

In this case study 2, we have presented an application of CABC framework and modeled the human behaviors in the form of complex network. The structural characteristics of the whole network have been studied and then the centrality measures of the strongly connected biggest cluster are calculated.

This case study is also an extension of the previous research in which centrality measures are validated over random networks and the question were raised that how these centrality measures perform in a real world. We apply same centrality measures over the real world network by collecting our own data. However, the resulting patterns of centrality metrics remained same in this network study [10].

By applying these measures and then after statistically analyzing the attribute data of the people under study, it is deduced that centrality measure results are not necessarily correlated with the real-world especially if the subjects are human.

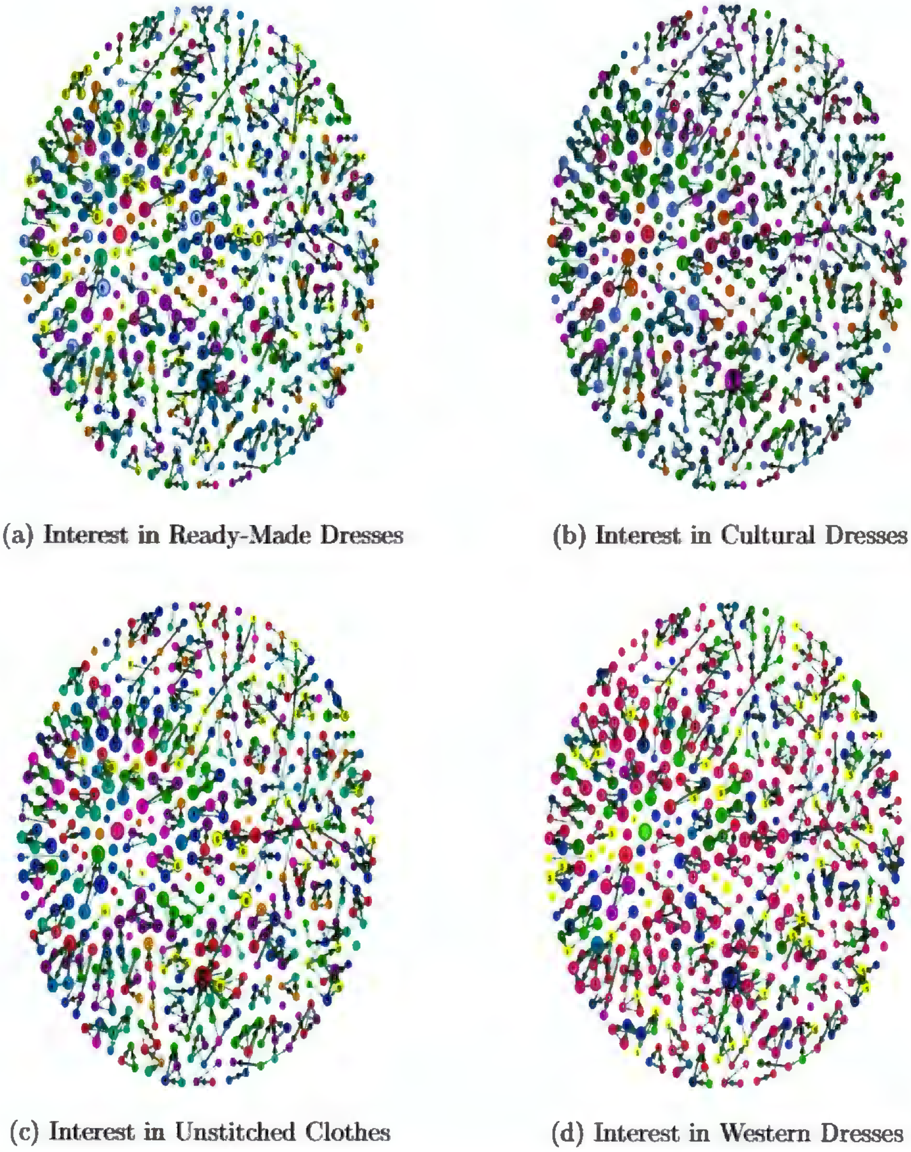
Conclusion This case study is considered as an example of complex system where people participating in a network have complex interactions with each other. In the end we didn't discover any emergent behavioral pattern as no behavioral influence is observed in the network of friends.

5.3 Comparison of Case Studies

After performing both case studies, we have applied both techniques used in case studies on the same data attributes and made comparison either which case study results give us detailed analysis.

For the case study 1, results have shown earlier, regarding their age 5 and interest in ready made dresses 6a, unstitched clothes 6c, cultural dresses 6b and western clothes 6d. When we have applied complex network modeling technique on the same data, we got more detailed results. Here we are showing the network models and the data description.

5.3.1 Visual Analysis of Interest in Different Dresses



(a) Interest in Ready-Made Dresses

(b) Interest in Cultural Dresses

(c) Interest in Unstitched Clothes

(d) Interest in Western Dresses

Figure 14: Network Models of Clothes Interests

Detailed Description of Interest in Ready-Made Dresses 14a

Node Color	Interest Rate	Percentage
Parrot Green	10	15.06%
Sea Green	8	14.46%
Blue	5	12.45%
Mustard	6	10.84%
Sky Blue	9	9.44%
Magenta	7	9.04%
Shocking Pink	3	8.03%
Brown	2	7.43%
Royal Blue	4	7.23%
Purple	1	5.82%

Detailed Description of Interest in Cultural Dresses 14b

Node Color	Interest Rate	Percentage
Green	2	14.26%
Light Purple	5	13.86%
Blue	10	13.05%
Light Green	8	11.45%
Magenta	9	9.24%
Shocking Pink	6	9.04%
Baby Pink	3	8.43%
Brown	7	7.43%
Dark Purple	4	7.03%
Royal Blue	1	6.020%

Detailed Results of Interest in buying Unstitched Clothes 14c

Node Color	Interest Rate	Percentage
Red	8	15.06%
Magenta	10	13.25%
Blue	5	13.05%
Parrot Green	9	12.45%
Aqua Green	7	12.05%
Pink	2	8.43%
Mustard	6	7.83%
Brown	1	6.22%
Blue	3	6.02%
Sky Blue	4	5.420%

Detailed Description of Interest in Western Dresses 14d

Node Color	Interest Rate	Percentage
Pink	1	41.97%
Blue	2	13.45%
Yellow	5	11.45%
Parrot Green	3	8.43%
Red	4	5.42%
Green	7	5.02%
Maroon	6	4.82%
Navy Blue	8	3.82%
Sky Blue	10	2.81%
Purple	9	2.61%

By getting these results, we concluded that complex network modeling gives us more detailed and interesting findings.

5.4 Comprehensive Discussion

The people who interact with each other in a non-linear way can be considered as an example of complex system.

In this research, we have modeled the behavior of people to infer behaviors in a social network system.We have used formal specification model in Z language. These specifications has been used to present clear and unambiguous underlying complex social system.

Two validation case studies are conducted as a proof of concept. Complex network modeling is performed to discover behavioral patterns. Additionally we have conducted a case study1 which have explored the age-related pattern in a complex social system.

In the first case study, we have explored the emergent behaviors of dressing interests of female students, who showed their loss of interest in different types of dressing as their age increases. Our study results have correlated with the previous studies of age and dressing interests. We have conducted this study by doing statistical analysis. These results have helped us in just understanding the general behavior of people.

Our second case study is based on complex network modeling. For this study we took a greater sample of the human data than the first case study, made a complex network and conducted social network analysis.

This methodological case study has offered the model of real world complex social system.

In this study, structural characteristics of the complex network are studied and the centrality measures are calculated. Statistical analysis is done to analyze the behavioral influence of girls under study over their friends located at the first degree of separation. No behavioral influences are discovered regarding their food choices, sleeping habits and physical activities (exercise) etc.

From this case study results (regarding SNA, statistical analysis, centrality based analysis) We can consider two main factors contributing towards these findings. (i) the one is people participating in this study belong from different backgrounds/ cultures and have strong family and cultural influence. Secondly they might don't spend years with each other to get influence. (ii) This is also deduced that centrality results are not necessarily correlate with the real world especially when the subject is human. Due to this dissimilar pattern of network centrality metrics and the real world behavior we have concluded that, we need to consider temporal data of the same cohort and apply the same techniques to again analyze their behavioral influences.

In the end, we have made comparison of both case study results and concluded that complex network modeling have provided us with the greater understanding of the social system and the detailed analysis of the behavior of people participating in a social network as compared to the old traditional techniques of statistical analysis. We have proved that complex network modeling is an effective way to analyze the choices/behaviors of people participating in a social system.

This study is also an extension of the previous research[10] on centrality measures. Centrality measures are analyzed in a real world human subject data which were demanded in the previous study. However, the resulting pattern of the centrality measures remained same in this network study.

6 Conclusion

The final chapter of the thesis concludes the research by discussing (i) the research question and the motivation behind the research (ii) the main contributions and findings (iii) and some additional findings as an extension of previous work.

In the past, frameworks such as cognitive agent based computing and formal agent based simulation have been introduce to study the behavior of complex systems. However cognitive agent based computing never have been used to study the influence of social choices.

In this thesis, we have presented a novel application of cognitive agent based computing framework for modeling human behavior. The main goal was to develop a hybrid model made of formal specification model coupled with complex network modeling which allowed us to analyze human behavior. Formal specifications have been written to model the behavior in a clear and unambiguous way. Then conducted two validation case studies. The first one is the case study survey and is based on statistical analysis, second case study is the application of complex network modeling level of cognitive agent based computing framework. In complex network modeling, social network analysis techniques are applied to evaluate the behavior of participants.

First of all we have analyzed the topological structure of the whole network. To analyze the key individuals of the network regarding their central positions, power, gregariousness, influential personality centrality based analysis is used. Then eventually we have analyzed the behavioral influences of the people over each other. This methodology presents an idea of modeling of human behavior with the developed hybrid model and also an application of complex network modeling of human behavior with the help of social network analysis.

Additionally in the end, we have also made comparison of two case studies to show the efficacy of complex network modeling technique for analyzing human behavior.

This research contributes as an extension of the previous research[10] on centrality measures. Centrality measures are analyzed in a real world human subject data which were demanded in the previous study. However, the resulting pattern of the centrality measures remained same in this network study.

6.1 Recommendations for Future Research

For future research, we recommend taking the temporal data of the same cohort to analyze the behavioral patterns and influences of people participating in the social network.

This will help in eliminating the doubts on our findings which gave us different results of the structural analysis of a network and behavioral analysis of the people which were expected to be the same.

References

[1] Lewis R Aiken Jr. The relationships of dress to selected measures of personality in undergraduate women. *The Journal of Social Psychology*, 59(1):119–128, 1963.

[2] Mir M Ali, Aliaksandr Amialchuk, and Silda Nikaj. Alcohol consumption and social network ties among adolescents: evidence from add health. *Addictive behaviors*, 39(5):918–922, 2014.

[3] Franck Amadieu, Claudette Mariné, and Carole Laimay. The attention-guiding effect and cognitive load in the comprehension of animations. *Computers in Human Behavior*, 27(1):36–40, 2011.

[4] Pavlos Antoniou and Andreas Pitsillides. Understanding complex systems: a communication networks perspective. Technical report, Technical Report TR-07-01, Department of Computer Science, University of Cyprus, 2007.

[5] Mariana Astudillo, Jennie Connor, Rachel E Roiblatt, Akanidomo KJ Ibanga, and Gerhard Gmel. Influence from friends to drink more or drink less: a cross-national comparison. *Addictive behaviors*, 38(11):2675–2682, 2013.

[6] Robert M Axelrod. *The complexity of cooperation: Agent-based models of competition and collaboration.* Princeton University Press, 1997.

[7] Yaneer Bar-Yam. *Dynamics of complex systems*, volume 213. Addison-Wesley Reading, MA, 1997.

[8] Yaneer Bar-Yam. A mathematical theory of strong emergence using multiscale variety. *Complexity*, 9(6):15–24, 2004.

[9] Mathieu Bastian, Sebastien Heymann, Mathieu Jacomy, et al. Gephi: an open source software for exploring and manipulating networks. *ICWSM*, 8:361–362, 2009.

[10] Komal Batool and Muaz A Niazi. Towards a methodology for validation of centrality measures in complex networks. *PloS one*, 9(4):e90283, 2014.

[11] Judith E Blakemore, Asenath A LaRue, and Anthony B++ Olejnik. Sex-appropriate toy preference and the ability to conceptualize toys as sex-role related. *Developmental Psychology*, 15(3):339, 1979.

[12] Deanna Blansky, Christina Kavanaugh, Cara Boothroyd, Brianna Benson, Julie Gallagher, John Endress, and Hiroki Sayama. Spread of academic success in a high school social network. *PloS one*, 8(2):e55944, 2013.

[13] Phillip Bonacich. Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology*, 2(1):113–120, 1972.

[14] Francesco Bonchi, Carlos Castillo, Aristides Gionis, and Alejandro Jaimes. Social network analysis and mining for business applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3):22, 2011.

[15] Stephen P Borgatti, Kathleen M Carley, and David Krackhardt. On the robustness of centrality measures under conditions of imperfect data. *Social networks*, 28(2):124–136, 2006.

[16] Florian Boudin. A comparison of centrality measures for graph-based keyphrase extraction. In *International Joint Conference on Natural Language Processing (IJCNLP)*, pages 834–838, 2013.

[17] Nicholas A Christakis and James H Fowler. The spread of obesity in a large social network over 32 years. *New England journal of medicine*, 357(4):370–379, 2007.

[18] Nicholas A Christakis and James H Fowler. The collective dynamics of smoking in a large social network. *New England journal of medicine*, 358(21):2249–2258, 2008.

[19] Lucy J Cooke and Jane Wardle. Age and gender differences in children's food preferences. *British Journal of Nutrition*, 93(05):741–746, 2005.

[20] Wouter de Nooy, Andrej Mrvar, and Vladimir Batagelj. *Exploratory social network analysis with Pajek*, volume 27. Cambridge University Press, 2005.

[21] Anne E Dickerson and Anne G Fisher. Age differences in functional performance. *American Journal of Occupational Therapy*, 47(8):686–692, 1993.

[22] Anne E Dickerson and Anne G Fisher. Effects of familiarity of task and choice on the functional performance of younger and older adults. *Psychology and aging*, 12(2):247, 1997.

[23] Maloa Ebeling and ML Rosencranz. Social and personal aspects of clothing for older women, 1961.

[24] R Darin Ellis and Jason C Allaire. Modeling computer interest in older adults: The role of age, education, computer knowledge, and computer anxiety. *Human factors: The journal of the human factors and ergonomics society*, 41(3):345–355, 1999.

[25] Joshua M Epstein. Why model? *Journal of Artificial Societies and Social Simulation*, 11(4):12, 2008.

[26] James H Fowler, Nicholas A Christakis, Steptoe, and Diez Roux. Dynamic spread of happiness in a large social network: longitudinal analysis of the framingham heart study social network. *BMJ: British medical journal*, pages 23–27, 2009.

[27] Linton C Freeman. Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239, 1979.

[28] Noah E Friedkin. Theoretical foundations for centrality measures. *American journal of Sociology*, pages 1478–1504, 1991.

[29] Ronald E Goldsmith, Leisa Reinecke Flynn, and Mary Ann Moore. The self-concept of fashion leaders. *Clothing and Textiles Research Journal*, 14(4):242–248, 1996.

[30] Mark S Granovetter. The strength of weak ties. *American journal of sociology*, pages 1360–1380, 1973.

[31] Per Hage and Frank Harary. Eccentricity and centrality in networks. *Social networks*, 17(1):57–63, 1995.

[32] Anthony Hall. Seven myths of formal methods. *Software, IEEE*, 7(5):11–19, 1990.

[33] Jeffrey A Hall and Thomas W Valente. Adolescent smoking networks: The effects of influence and selection on future smoking. *Addictive behaviors*, 32(12):3054–3059, 2007.

[34] Robert A Hanneman and Mark Riddle. Introduction to social network methods, 2005.

[35] Robert A Hanneman and Mark Riddle. Introduction to social network methods, 2005.

[36] George W Hartman. Clothing: personal problem and social issue. *Journal of Home Economics*, 41(6):295–298, 1949.

[37] Alison L Hill, David G Rand, Martin A Nowak, and Nicholas A Christakis. Emotions as infectious diseases in a large social network: the sisa model. *Proceedings of the Royal Society B: Biological Sciences*, 277(1701):3827–3835, 2010.

[38] Ben T Hirsch, Suzanne Prange, Stephanie A Hauver, and Stanley D Gehrt. Raccoon social networks and the potential for disease transmission. *PloS one*, 8(10):e75830, 2013.

[39] Bill Howard. Analyzing online social networks. *Communications of the ACM*, 51(11):14–16, 2008.

[40] Michael J Jacobson and Uri Wilensky. Complex systems in education: Scientific and educational importance and implications for the learning sciences. *The Journal of the learning sciences*, 15(1):11–34, 2006.

[41] Elihu Katz and Paul Felix Lazarsfeld. *Personal Influence, The part played by people in the flow of mass communications*. Transaction Publishers, 1970.

[42] Arpita Khare, Ankita Mishra, Ceeba Parveen, and Rajlaxmi Srivastava. Influence of consumers susceptibility to interpersonal influence, collective self-esteem and age on fashion clothing involvement: A study on indian consumers. *Journal of Targeting, Measurement and Analysis for Marketing*, 19(3):227–242, 2011.

[43] Eric D Kolaczyk and Gábor Csárdi. *Statistical analysis of network data with R*, volume 65. Springer, 2014.

[44] Joy Michelle Kozar. Effects of model age on adult female consumers purchase intentions and attitudes for an age-specific product, clothing. *International Journal of Marketing Studies*, 4(2):p22, 2012.

[45] Jinkook Lee, Sherman D Hanna, Chiu Fui Joyce Mok, and Hui Wang. Apparel expenditure patterns of elderly consumers: A life-cycle consumption model. *Family and Consumer Sciences Research Journal*, 26(2):109–140, 1997.

[46] Naoki Masuda, Issei Kurahashi, and Hiroko Onari. Suicide ideation of individuals in online social networks. 2013.

[47] Sara C Mednick, Nicholas A Christakis, and James H Fowler. The spread of sleep loss influences drug use in adolescent social networks. *PloS one*, 5(3):e9775, 2010.

[48] Melanie Mitchell. *Complexity: A guided tour*. Oxford University Press, 2009.

[49] Mark EJ Newman. The structure and function of complex networks. *SIAM review*, 45(2):167–256, 2003.

[50] Muaz A Niazi. *Cognitive Agent-based Computing-I: A Unified Framework for Modeling Complex Adaptive Systems Using Agent-based & Complex Network-based Methods*, volume 1. Springer, 2012.

[51] Muaz A Niazi. Emergence of a snake-like structure in mobile distributed agents: An exploratory agent-based modeling approach. *The Scientific World Journal*, 2014, 2014.

[52] Muaz AK Niazi. Towards a novel unified framework for developing formal, network and validated agent-based simulation models of complex adaptive systems. 2011.

[53] Muaz AK Niazi, Amir Hussain, and Mario Kolberg. Verification and validation of agent based simulations using the vomas (virtual overlay multi-agent system) approach. 2009.

[54] Marion O'Brien and Aletha C Huston. Development of sex-typed play behavior in toddlers. *Developmental Psychology*, 21(5):866, 1985.

[55] Aron O'Cass. Fashion clothing consumption: antecedents and consequences of fashion clothing involvement. *European Journal of Marketing*, 38(7):869–882, 2004.

[56] Aron O'Cass and Eric Choy. Studying chinese generation y consumers' involvement in fashion clothing and perceived brand status. *Journal of Product & Brand Management*, 17(5):341–352, 2008.

[57] Tore Opsahl, Filip Agneessens, and John Skvoretz. Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, 32(3):245–251, 2010.

[58] Clyde C Robinson and James T Morris. The gender-stereotyped nature of christmas toys received by 36-, 48-, and 60-month-old children: A comparison between nonrequested vs requested toys. *Sex Roles*, 15(1-2):21–32, 1986.

[59] Shahar Ronen, Bruno Gonçalves, Kevin Z Hu, Alessandro Vespignani, Steven Pinker, and César A Hidalgo. Links that speak: The global language network and its association with global fame. *Proceedings of the National Academy of Sciences*, 111(52):E5616–E5622, 2014.

[60] Mary Lou Rosencranz. *The application of a projective technique for analyzing clothing awareness, clothing symbols, and the range of themes associated with clothing behavior*. PhD thesis, Michigan State University of Agriculture and Applied Science. Department of Sociology and Anthropology, 1960.

[61] Mary Lou Rosencranz. *Clothing concepts: A social-psychological approach*. Macmillan New York, 1972.

[62] J Niels Rosenquist, Joanne Murabito, James H Fowler, and Nicholas A Christakis. The spread of alcohol consumption behavior in a large social network. *Annals of Internal Medicine*, 152(7):426–433, 2010.

[63] Richard B Rothenberg, John J Potterat, Donald E Woodhouse, William W Darrow, Stephen Q Muth, and Alden S Klovdahl. Choosing a centrality measure: epidemiologic correlates in the colorado springs study of social networks. *Social Networks*, 17(3):273–297, 1995.

[64] Mary S Ryan. Psychological effects of clothing. 1954.

[65] Giovanni Scardoni and Carlo Laudanna. *Centralities based analysis of complex networks.* INTECH Open Access Publisher, 2012.

[66] Anna Servin, Gunilla Bohlin, and Lisa Berlin. Sex differences in 1-, 3-, and 5-year-olds' toy-choice in a structured play-session. *Scandinavian Journal of Psychology*, 40(1):43–48, 1999.

[67] Carl Shapiro and Hal R Varian. *Information rules: a strategic guide to the network economy.* Harvard Business Press, 2013.

[68] RS Sharma. Clothing behaviour, personality, and values: A correlational study. *Psychological studies*, 1980.

[69] S Shipley and ML Rosencranz. Older womens clothing preferences. *Journal of Home Economics*, 54(10):854–854, 1962.

[70] Amnah Siddiqa and Muaz Niazi. A novel formal agent-based simulation modeling framework of an aids complex adaptive system. *International Journal of Agent Technologies and Systems (IJATS)*, 5(3):33–53, 2013.

[71] Steven H Strogatz. Exploring complex networks. *Nature*, 410(6825):268–276, 2001.

[72] ME Tyrchniewicz and CA Gonzales. relationship between specific clothing variables and self concept of adult women. *Canadian home economics journal*, 1978.

[73] Bert N Uchino, Timothy W Smith, and Cynthia A Berg. Spousal relationship quality and cardiovascular risk dyadic perceptions of relationship ambivalence are associated with coronary-artery calcification. *Psychological science*, 25(4):1037–1042, 2014.

[74] Jennifer B Unger, Michele D Kipke, Christine J De Rosa, Justeen Hyde, Anamara Ritt-Olson, and Susanne Montgomery. Needle-sharing among young iv drug users and their social network members: The influence of the injection partner's characteristics on hiv risk behavior. *Addictive behaviors*, 31(9):1607–1618, 2006.

[75] Stanley Wasserman and Katherine Faust. *Social network analysis: Methods and applications*, volume 8. Cambridge university press, 1994.

[76] Duncan J Watts. *Small worlds: the dynamics of networks between order and randomness.* Princeton university press, 1999.

[77] Jim Woodcock, Peter Gorm Larsen, Juan Bicarregui, and John Fitzgerald. Formal methods: Practice and experience. *ACM Computing Surveys (CSUR)*, 41(4):19, 2009.

Appendices

Questionnaire (Case Study 1)

Analysis of Girls Hostel

1. Name

____

2. Age

____

3. Give name of 3 Friends inside the hostel

a.______________

b.______________

c.______________

4. Number of hours of daily computer use

a. 1-3

b. 4-6

c. 6+

5. Number of mobile sims that u use

____

6. Number of contacts in your mobile

____

7. Health status

a. Fit

b. Fat

c. Weak

d. Average

8. Interest in buying western dresses: (1-5) (1: low , 5: max)

____

9. Interest in buying cultural dresses: (1-5)

____

10. Interest in buying ready made dresses: (1-5)

____

11. Interest in buying unstitched clothes (1-5)

____

12. Rate 3 cloth brands according to your choice

a.————-

b.————-

c.—————-

13. Where do you usually take maximum inspiration for new cloth purchase(name them according to your preferences)

a. Friends

b. Family

c. Celebrities

d. Magazines

e. Style in stores

f. Other

14. In every year how often do you purchase a new item of clothing:

____

15. What would be the period of year you spend most on clothing

a. Spring

b. Summer

c. Autumn

d. Winter

16. What is the percentage of clothing and accessories that you purchase are bought online

a. 0 %

b. Less than 50 %

c. More than 50 %

17. Name the following in terms of your priority when purchasing clothes

a. Qualities

b. Price

c. Comfort

d. Style

e. Color

f. Material

g. Brand

18. Who do you mainly shop with

a. Alone

b. Friends

c. Family

19. When purchasing a new cloth item, do you buy a matching outfit item (shoes, bag, coat)

a. No

b. Not very often

c. Sometimes

d. All the time

Questionnaire (Case Study 2)

Analysis of Girls Hostel

1. Name (Please write your full name)

____

2. Age

____

3. Give name of 3 friends inside the hostel: (a=best friend, b= 2nd best friend, c= 3rd best friend)
(Please write their full names)
a. ______________
b. ______________
c. ______________
4. How often do you meet your 1st best friend?
a. Once a day
b. Twice a day
c. Almost all the time
5. How often do you meet your 2nd best friend?
a. Once a day
b. Twice a day
c. Almost all the time
6. How often do you meet your 3rd best friend?
a. Once a day
b. Twice a day
c. Almost all the time
7. On a scale of 1 to 10, number of hours of daily computer use

____

8. Number of mobile Sims that you use

____

9. Number of contacts in your mobile

____

Dressing Interests

10. On a scale of 1 to 10, interest in buying western dresses: (Min: 1, Max: 10)

____

11. On a scale of 1 to 10, interest in buying cultural dresses

____

12. On a scale of 1 to 10, interest in buying Readymade dresses

____

13. On a scale of 1 to 10, interest in buying unstitched clothes

____

14. On a scale of 1 to 10, where do you usually take maximum inspiration for new cloth purchase? (Please rate all the options)

a. Friends ———-

b. Family ———

c. Celebrities ———

d. Magazines ———

e. Style in stores ————

f. Other———-

15. In every year how often do you purchase a new item of clothing?

a. 5 times

b. 12 times

c. 15 times

d. More than 15 times.

16. What would be the period of year you spend most on clothing? (On a scale of 1 to 4; rate the following) (4: Max, 1: Min)

a. Spring ———

b. Summer———

c. Autumn ———

d. Winter———

17. Do you purchase cloth and accessories online?

a. Yes

b. No

18. On a scale of 1 to 10, rate the following in terms of your priority when purchasing clothes

a. Qualities ———

b. Price ———

c. Comfort ———

d. Style ———

e. Color ———

f. Material ———

g. Brand———

19. Do you prefer dark colored clothes?

a. Yes

b. No

20. Do you prefer light colored clothes?

a. Yes

b. No

21. Who do you mainly shop with?

a. Alone

b. Friends

c. Family

22. When purchasing a new cloth item, do you buy a matching outfit item (shoes, bag, and coat?)

a. Yes

b. No

23. Rate 3 cloth brands according to your choice (1: Max, 3: Min)

1.————-

2.————-

3.————-

4. Don't wear brands

Health Data

24. How do you perceive your health?

a. Fit

b. Fat

c. Weak

d. Average

25. On a scale of 1 to 10, how active are you?

____

26. What is the frequency of your activity in a week: (Sports or Exercise Etc?)

a. Don't exercise

b. 1 day/week

c. 2 days

d. 3 days

e. 4 days

f. 5 days

g. 6 days

h. 7 days

27. On a scale of 1 to 10, how much do you like fast food?

____

28. On a scale of 1 to 10, how much do you like home cooked?

____

29. On average, how many hours do you sleep?

____

30. On a scale of 1 to 10, how often do you take fast food in a week?

____

31. On a scale of 1 to 10, how often do you take carbonated drinks in a week?

____

32. On a scale of 1 to 10, how often do you take packed juices in a week?

____

33. Do you smoke?

a. Yes

b. No

34. How often do you smoke?

a. Everyday
b. Once a week
c. Less than once a week
d. Don't smoke

35. How often do you take vegetables?
a. Never
b. Less than once a week
c. Once a week
d. More than once a week
e. Once daily
f. More than once daily

36. How often do you take fruits/fresh juices?
a. Never
b. Less than once a week
c. Once a week
d. More than once a week
e. Once daily
f. More than once daily

37. How often do you take chocolates/ice creams?
a. Never
b. Less than once a week
c. Once a week
d. More than once a week
e. Once daily
f. More than once daily