Studying Learning Networks within Moodle: A Social Network Analysis Approach

vorgelegt von

M. Sc. Rania Al Halaseh

aus Madaba, Jordanien

Von der Fakultät II
Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften
Dr. rer. nat.

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. Christian Thomsen
Gutachter: Prof. Dr. Lars Knipping
Gutachter: Prof. Dr. Wilfried Hendricks
Gutachterin: Prof. Dr. Sabina Jeschke

Tag der wissenschaftlichen Aussprache: 22.Mai 2014

Berlin 2014

Contents

List of Figures

List of Tables

Abbreviations

ALN	Asynchronous Learning Network
CoP	Communities of Practice
CAI	Computer Assisted Instruction
CBI	Computer Based Training
CSCL	Computer Supportive Collaborative Learning
CMC	Computer Mediated Communication
DM	Data Mining
DLMS	Distance Learning Management System
DIT	Diffusion of Innovations Theory
EDM	Educational Data Mining
FOAF	Friend of a Friend
GUI	Graphical User Interface
KDD	Knowledge Data Discovery
ICT	Information Communication Technology
ISIS	Information System for Instructors and Students
IS	Information System
LMS	Learning Management System
LA	Learning Analytics
LLL	Lifelong Learning
MOOC	Massive Open Online Course
Moodle	Modular Object Oriented Developmental Learning Environment
PLEF	Personal Learning Environment Framework
PLN	Personal Learning Network
PLENK	Personal Learning Environments, Networks and Knowledge

SNA Social Network Analysis

SCORM Sharable Content Object Reference Model

SNAPP Social Networks Adapting Pedagogical Practice

STD Standard Deviation

SNS Social Network Site

SQL Structured Query Language

TEL Technology Enhanced Learning

VLE Virtual Learning Environment

Acknowledgment

First of all, I would like to express my sincere and deepest gratitude for my professor Prof. Lars Knipping for everything he did for me. I will always be thankful for him for accepting me as his PhD student. His continuous support and guidance were very much helpful throughout my research.

I would like to thank Mr. Erhard Zorn for his continuous help and understanding, his door was always open for me. He always made sure that I'm welcomed as a part of the group.

I would like to thank Vera Kockler, my first German friend. Vera taught me everything about living here. You are just the best my friend.

Thank you Martina Blessing, Rafael, and Lilly for everything you did to me. You all were my family in Berlin.

Special thanks to my colleague Christian Grümme. I do appreciate your continuous encouragement and support. Thank you my friend for everything.

My great friends Michael, Francisco, Marvin, Maria Fernanda, Mridula, Helena, Uli, Eva, Monika, Helmut and the rest of KHG group in Kassel. I always had a great time with you all and made me forget all the problems and the struggles. Thank you for all laughters.

Finally, to all my friends just being in my life made Germany as a second home for me. You are all the only good thing that happened to me from this experience. You never pushed me to speak the language. You made it clear that it is not about the language we talk but it is more important that we talk and connect. Every and each one of you has a special place in my heart.

Dedication

"I can do everything through Christ who strengthens me." (Philippians 4:13)

"God has a reason for allowing things to happen. We may never understand His wisdom, but we simple have to trust His will." (Psalm 37:5)

I warmly dedicate this thesis to the most important people in my life, my parents, Azmi and Bouthayna Al Halaseh. Without them being by my side my whole life, I would never achieve this. Their unconditional love, continuous encouragement, and believing in me made my thesis come true.

Mama my best friend and Baba my hero, you are the best parents anyone can have. Nothing I can say or do will express how much I love you and appreciate everything you both did and still doing for me and my sisters and brother. Thank you for being by parents. Love you Mama and Baba.

I would like to thank my sweet wonderful sisters Rana and Rand, and my lovely brother Khaleel. Thank you for being so patient, supportive, and encouraging. Thank you for everything you did to cheer me up. You are the best sisters and brother ever. You know how much I love you all.

Abstract

This thesis presents a case study on analyzing students' participation level within modules from a Learning Management System (LMS) by inspecting the discussion forums and wikis. It studies levels of participation of students as well as the patterns of interactions formed by students, teachers, and tutors. These are investigated from a Social Network Analysis (SNA) viewpoint. In spite of its popularity in research areas like sociology, SNA has been only recently applied to study learning environments.

Most of the courses by higher learning institutions use LMSs. Basically, these systems serve as a Content Management System (CMS) to distribute course materials among students and have them doing assignments and quizzes. LMSs provide a number of communication services such as discussion forums, wikis, messages, and chats. These can enhance the interactive social nature of learning by communication and collaboration between students.

This thesis provides a quantitative approach for studying learning networks extracted from discussion forums and wikis. The basic approach is to represent these interactions as a network where nodes represent either students or teachers and links the relation taking place between them. SNA is used as an analytical tool to study the characteristics of the relations within a network and in which ways they reflect learning process in general within the course, e.g. the level of participation.

The results obtained from analyzing learning networks provide an insight to the course practices. They provide feedback to teachers about the different roles students occupy within the course along with its relation to their educational performance. Studying the formation of groups within the course can improve the quality of knowledge building, information distribution among the course, and foster their collaboration. As a conclusion, learning networks in a formal context in LMS are as important as these in an informal one like Facebook. Thus, giving larger attention for studying formal networks can enrich students' collaboration and problem solving levels within the course.

Chapter 1

Introduction

1.1 Introduction

Humans in general prefer to work jointly with others, either friends, family members, or colleagues, to communicate, share ideas, and work jointly on problems. These connections, interactions, and relationships can be studied and analyzed from social network perspective. The large amounts of information that can be extracted from these social interactions can be very useful to get an overall understanding of the studied phenomena. These extracted network represent different types of relationships formulated between different individuals such as friends, colleagues, classmates, roommates, etc. Considering the large developments in technological applications and the uses of Internet applications in almost every single aspect of daily life i.e. education, shopping, electronic governments, the is huge amounts of online data about user's individuals and their personal information. These online data which represent the interactions in online applications along with personal information are represented in social networks and later on understand and analyze the relations formed within them. This concept was emphasized through this idea: "Computer networks are inherently social networks, linking people, organizations, and knowledge". Internet and web applications serve as virtual environments where individuals meet and exchange ideas (Wellman, 2001).

Consequently, new research approaches, concept, and terminology have been developed to benefit from huge amounts of online data and provide recommendations and advice for individuals. The social networks can be extracted from online applications like forums, chats, e-mails, wikis, blogs, and even from online social networks for social encounters. These online applications are the result of web developments which is marked with the introduction of *Web 2.0* applications. Web 2.0 applications are almost everywhere and used by all the people without any restriction or requiring previous education (Downes, 2005).

Learning is a process that requires communication, interaction, and participation of learners with their educators and the learning materials. Acquiring new knowledge is seen as a collaborative process which is maximized by working in groups. These

groups are formed by students, teacher, and tutors along with their social interactions in one common research direction. Formally, students tend to learn in study focus groups formed from their classmates where they meet together face-to-face and solve their problems and homeworks together. However, these face-to-face encounters and study groups have been replace by different online applications and Web 2.0 technologies. Specifically, Web 2.0 applications and technologies have a huge influence on educational approach in almost every aspect of learning either formal or informal approaches. The combination and collaboration between of Web 2.0 technologies in learning aspects is referred to as *E-learning 2.0* (Ebner, 2007).

Recently, the new hype in learning methods is engaging Web 2.0 applications into classroom and institutional teaching applications. This suggests integrating these online applications into traditional classroom and formal institutional education. Moreover, these applications can enhance learning through informal applications. These new learning aspects have many advantages on the learning process such as introducing newer methods for learning, solving problems, and knowledge acquiring, other than being focused on teachers only to get the information (Downes, 2005).

Nowadays, hybrid learning approach is the most highly used approach. Hybrid learning is represented as a supplement of classroom learning approaches, which is enhancing face-to-face teaching methods by additional online methods i.e. quiz, homework, presentations. All of that results in endless online applications and larger amounts of online data that can be analyzed to understand the learning environments or platforms. One of the different analytical methods is studying these students information and interactions by representing them as a social network and later on analyzing these networks' structures and patterns (Power *et al.*, 2010).

There are many used terms that describe the idea of e-learning, the different terms describe different levels of technology use in learning practices. These terms vary from totally online courses i.e. distance education, to hybrid online learning i.e. using online environments for aiding classroom teaching practices. In literature, it is defined the basic e-learning aspects, they are learning objects, learning communities, and social educational network services. Basically, learning objects represents older learning approaches, referred to as e-learning 1.0 whereas social networks along with learning communities represent newer learning environments and practices called e-learning 2.0 (Carugati *et al.*, 2008).

However, the oldest online learning aspects which are marked by e-learning 1.0 are referred to as formal learning aspects. The reason for this naming is that this type of learning is taking place under university or institution boundaries and is restricted to exams and grading systems. These formal learning approaches are initiated by teachers and tutor and the communications and interactions are supervised by them, on the other hand, students' role is restricted to download learning materials i.e. PowerPoint presentations, submitting assignments, taking exams, or participating

in online discussions with other classmates i.e. online forums of the class. On the other hand, e-learning 2.0 approaches and online courses are referred to as informal learning practices. This is because these learning approaches are not restricted to be held in classroom and any student can register in these applications and services regardless being registered in learning institution or not. These informal learning approaches allow learners to set their personalized learning and establish their own environments (Attwell, 2007a).

The introduction of new learning environments and practices are imposing a lot of pressure on formal learning approaches and their continuity. Students are more familiar and comfortable in using online services for understanding new concepts and finding solutions for their problems. However, learning institutions have introduced some of the informal services into their learning especially collaborative services applications to encourage students group formation. More specifically, Web 2.0 services and tools are introduced into blended learning approaches such as combining institutional learning environments i.e. face-to-face learning with Web 2.0 tools that support collaborative and cooperative group learning (Attwell, 2007a).

This thesis examines applying social network analysis approaches and measures to study and examine learning related problems. Social network theories studies the different types of social relations for any problem as long as it is represented as a social networks. Using social network theories in learning environments and problems have many advantages such as understanding how learning is taking place among these environments, finding out the different roles of students, and how new ideas can be dismissed among members. It also investigates what are the criteria that govern group formulation from student and measures the strength of the relations among students within learning groups (Scott, 2000; Dawson, 2006).

A social network is a set of individuals and the relationships among them where the individuals are represented as nodes and the relationships are represented as connections or links among them. A social network is studied and analyzed using different measures that allow understanding the social pattern of connections among the individuals along with the overall characteristics of the network due to the structure of nodes and their connections. Studying the learning environments from a social network perspective allow understanding how the information/news are exchanged and distributed among classmates (Wasserman & Faust, 1994; Reffay & Chanier, 2002).

Moreover, studying learning environments from a social network perspective improves communication between students themselves and their teachers. This arises many questions for further investigations, such as: what are the types of social relationships that could be established in any learning environment? And how can studying these relations provide newer choices and paths for exchanging news/messages/information? Or what are the criteria that triggers groups formation and student collaboration within learning environment? All of these questions form a new research area that can take advantage of the large volumes of the data being produced either from

Learning Management System (LMS) within learning institutions (Haythornthwaite, 1996; Cho *et al.*, 2007; Dawson, 2006).

Consequently, social network analysis approach allow understanding the dynamics between students in the classroom or environment through studying the extracted social networks. The idea is to represent the social network of specific courses within the LMS where the course is represented as a social network of students and their connections. This could answer questions such as: how active are students within the course? How students tend to work collaboratively in smaller groups i.e. communities? The social network approach provides a tool to extract useful and meaningful information from large volumes of data produced from students' interactions within the LMS. After extracting social networks from LMS data, they can be later analyzed using social network theories (Scott, 2000; Garton *et al.*, 1997).

This thesis focuses on studying the social relations in courses within LMS. More specifically studying students' participation level within courses and the interaction levels among learning activities within the courses. Studying and analyzing participation level within learning activities influence many aspects in learning process such as learning outcomes, knowledge acquiring levels among students in the course. Social network theories studies the social networks among different participation levels among the course such as frequency of interactions among students, number of messages sent and received, and discovering who controls the communication in the courses, and finding out the outcomes of the students (Haythornthwaite, 2005; Scott, 2000).

1.2 Relation between LMSs and Social Network Analysis

E-learning is defined as the learning approach that combine technological applications online services into educational practices and approaches, in other words it defines the integration of Information Communication Technologies (ICT) into learning activities and methodology. A LMS defines and organize courses within learning institutions i.e. universities or schools. Moreover, LMS organizes and controls student's and teacher's personal and academic information as well as their interactions with their peers and teachers. Students interact within these environments through downloading learning materials and delivering homework. Also, LMS offer some online activities for encouraging collaborative learning approaches through their defined modules like discussion forums, wikis, messages which are all introduced among course boundaries and used by registered students (Haythornthwaite, 2005).

In comparison to face-to-face traditional classroom learning approaches, online learning environments produce large amounts of online data about the different interactions within them. These online learning systems have a large databases of

the information and data about the enrolled students, teachers, and tutors in the learning institution. These personal information i.e. students personal profiles, data, and interactions are save into log files in relational databases. These aggregated data from LMS's databases are very important and valuable for understanding the social relations among courses. Moreover, analyzing these information can provide a valuable review of student's participation levels, outcomes, learning styles within the course (Dawson, 2006; Haythornthwaite, 1998; Jones & Dirckinck-Holmfield, 2009).

However, large amounts of information (i.e. users profiles and their participation and academic data) saved in LMS need to be further analyzed to gain an overall understanding of course dynamics. Basically, the LMS have some basic summarizing tools for students information but they are only basic statistics like number of times a page was visited, downloaded materials, or modules visited (William, 2006).

This study uses Social Network Analysis (SNA) as an analytical tool to analyze and understand students interactions within LMS along with all the social aspects within LMS modules. Using SNA for examining social aspects within LMS provide a feedback about group characteristics, participation frequencies and levels, group formation between students, and social relationships between students themselves and their teachers within courses. Basically, applying SNA approach is based on representing interactions within the course as a social network. This study investigates the interactions resulting from discussion forums and wiki entries within modules of the LMS (Haythornthwaite, 2005).

The SNA approach is valuable for studying relationships among individuals. It is used as an analytical tool for studying the structures of the social networks, the different roles of students within the network due to their position within the network. Also, it allow understanding social networks focused on many analytical levels individual, group, and overall network level (Garton *et al.*, 1997).

1.3 Importance of Social Network Analysis

SNA is the study and analysis of social relationships among individuals and the social structure of their interactions once they are represented as a social networks. A social network is a set individuals and the relationships found between them. These nodes can represent individuals, students, employees, friends, family members, etc. Whereas the links among them represent relationships in organizations, family tree, schools, business, etc. There are two basic relationships that can be formed by social networks, first is one way relation i.e. friendship relation, second is membership relation i.e. an individual attending an event (Wasserman & Faust, 1994; Scott, 2000).

Originally, SNA as an analytical approach is used for studying social relations among individuals from various perspectives individual's level, subgroup levels, and the whole network level. Based on social network theory, individual's actions and performance are affected by their position within the network i.e. the individuals who are the immediate neighbors. Therefore, understanding how individuals behave within their environment, their relations to others, and the social structure can be done through analyzing the underlying social networks that represent them. All of these aspects have been studied by many researchers (Haythornthwaite, 1996, 1998; Wasserman & Faust, 1994; Wellman, 1996).

Using SNA is not only restricted to be used in social science researches for understanding individual's behavior or characteristics and how they interact with others. On the contrary, SNA have been applied to a wide range of problems or phenomena as long as their interactions can be represented as a social network. This is due to the large number of measures and theories that can be used for studying different aspect of the network. As examples of the various research approaches where SNA have been applied are studying obesity (Christakis & Fowler, 2007), terrorism (Krebs, 2002), business organizational relations (Burt, 2004; Gulati, 1995), knowledge management and building (Cohen & Prusak, 2001), and Internet applications (Wellman, 1996; Shane Dawson & Poole, 2008).

Recently, many studies have discussed applying SNA into learning contexts. For example, a study by Haythornthwaite (Haythornthwaite, 1999) used SNA for analyzing the quantities and types of relations and interactions established among distance learners. This study noticed the importance of forming groups and collaborative group activities among learners for maintaining stronger relationships among learners. This will foster the collaboration among learners and improve their outcomes.

There are many advantages of applying SNA into learning problems, especially through utilizing different SNA measures like density and centrality measures. A study by (Haythornthwaite, 2001) applied these measures for studying social networks of class interactions. This study defined many social network based on the relations among the class like collaborative group network, advice network among students, support network, and informal relations. This study intended to find answer of the following question: "Do the four relations describe similar structures or do they capture different aspects of student interaction?". In their findings, they found out the densities of these networks increased by the end of the semester and they work more with their group members. This was true also for advice network, since by time the trust increased between team members.

Another study, have studied the importance of centrality and cohesion measures in social networks to estimate collaborative learning level within a course. In their study, they defined a social network that represent the interactions within discussion forums within a course. They found out the densities among these networks were very low. The centrality measures pointed out that there are only two learners i.e. having

high centrality values, were dominating the interactions within the discussion forums. This limits interactions among larger number for students within the class (Reyes & Tchounikine, 2005).

The most recent work by (Cho *et al.*, 2007) investigated students' characteristics and role among the social network in the course. They utilized centrality measures i.e. degree, betweenness, and closeness and structural holes detection for studying relationships in the network. These measures studied the roles of students according to the position they occupy within the network. Most importantly they tried to correlate between the results of these measure and the students' success level.

These previous studies show that SNA is an important research approach for learning problems. It's measures can be used for studying different students characteristics, group structures, and standards for creating learning communities. Understanding the social dynamics between students, teachers, and the formed groups are considered as an important feedback for teachers about participation levels within the course and any clue about student's academic results.

1.4 Research Questions

This thesis studied the use of SNA, as an analytical tool, to investigate and analyze the different interactions within learning modules offered from LMS as learning environments. As explained before, a LMS produces large amounts of data that is the result of the students' participation within these environments. This study also employs SNA measures to investigate the different characteristics of the social patterns formed within the learning modules.

More specifically, this study analyzes the level of participation within two main collaborative modules within Moodle; discussion forums and wikis. These interactions form different social networks and can be subject for analysis by SNA tools. The result of this analysis can give an indication on how active the students are within the course's activities.

It also studies the different aspects of the Web 2.0 and its applications and effects on the learning process, as well as its learning theories and effect on the development of a new theory of connectivism. Moreover, it also focuses on the different personalized systems that are the result of integrating Web 2.0 technologies onto the LMS practices.

The following are the main research questions that are investigated throughout this study. They represent the main research focus:

- How useful is SNA as an analytical tool for studying relations within learning environments?

- What are the types of social learning networks formed in ISIS?
- What information can be obtained from studying these different network?
- How often were student groups built? And what is the basis for creating these networks?
- What are the characteristics of the social structure formed from wikis and discussion forums networks within ISIS?

One main drawback or limitation in using SNA as an analytical tool within learning environments and their interactions is the teacher's or administrator's concerns and worries about data privacy of the students. Since they are dealing with the students' grades and information (Wasserman & Faust, 1994).

2 Chapter 2 E-Learning 2.0 and Social Learning

2.1 Introduction

Many different terms and concepts have been used to illustrate the concept of using technology within education. All these concepts have a common objective of improving knowledge building and transfer. Two largely used terms are e-learning and Technology Enhanced Learning (TEL). They both describe the integration between technology and educational and learning methods. These two terms are used interchangeably (Dawson *et al.*, 2008).

E-learning encloses all the educational scenarios based on the applications of Information Communication Technology (ICT). One definition of e-learning is to describe a number of technologies and methods that can be applied to deliver courses and exercises in an electronic format. Many terms within this research field are to characterize this approach, for example *online learning*, *virtual learning*, *distributed learning*, and *web-based learning* (Gerhard *et al.*, 2002).

TEL however encloses a wider range of concepts and cases. TEL express the use of technology and their applications into learning uses. By definition, TEL contains all the technological services used to aid class teaching along with online websites that aids learning in informal ways. The difference between e-learning and TEL is that TEL activities and applications support learning in informal fashion without correlation with outcomes and grades (Schmees, 2006).

For several years, web-based applications developed for educational purposes have been associated with didactical evolutions. They are highly related to social interactions and communications between learning participants. All of that is a result of the large development in Web 2.0 applications and technologies (Schmees, 2006).

TEL and e-learning applications can be seen as a substitute to the traditional learning approaches i.e. VLE or LMS. The traditional approaches are based on delivering

learning material for students. On the contrary, the latest learning approaches focus on instant sharing and communicating with learning participants, and especially open access of learning material. The early examples of online web application are characterized by universities websites, digital libraries, and CMS. Afterwards, these online applications have witnessed the introduction of participative interactive systems. For example, these systems include: sending and receiving messages, live conferences, and online quizzes (Schmees, 2006).

The new definition of the Web, marked by the introduction of Web 2.0 applications, has a great effect on learning and educational uses. Web 2.0 has integrated social software applications, i.e. blogs, wikis, and online sharing applications into learning environments. This encourages the collaboration, cooperation, and interaction among learners. Since 2004, Web 2.0 has brought a new perception of the use of web, which highly empowers learners through giving them more possibilities of online services. This changed the role of learners and stopped being passive, only receiving learning material, but being more active and creators for learning material that are compatible with their own learning styles (O'Reilly, 2005).

As stated previously, these technological development have greatly affecting teaching approaches. Especially, these new approaches represent new relation between teachers and their students, characterized by high interactions and sharing ideas. In the literature, several learning approaches were defined, including: traditional learning (i.e. face-to-face), e-learning, blended learning, mobile learning, and personalized learning (Bates, 2010).

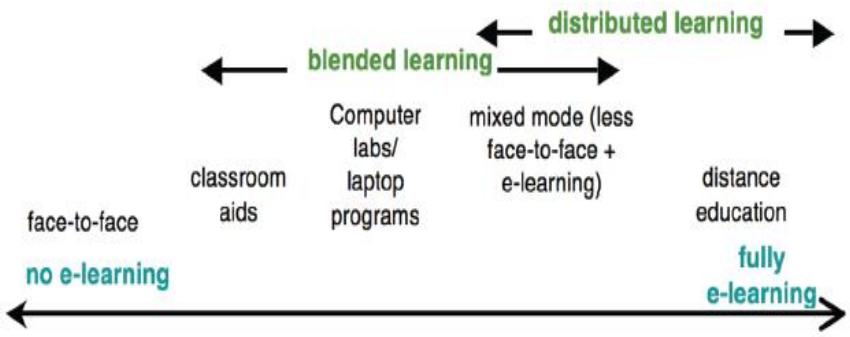


Figure 2.1: Different forms of e-learning (Bates, 2010)

Figure 2.1 shows the development in learning approaches from simple environments i.e. no technological applications, into completely online learning ones. Simple learning approaches represent classroom teaching where teachers distribute learning materials into a mass of students. In this case, the teacher have a full control of the learning plan and material structure and the distribution to all students. The learning material is distributed among students regardless their own learning styles and with minimal role for students (Bates, 2010).

One of the earliest cases of using technological applications in learning applications went back to 1980s which marked the beginning of Computer Assisted Instruction (CAI), which is called Computer-based training (CBI) as well. Their functionality include creating learning programs which teaches students using computers. The importance of CAI reside on providing teaching programs along with practicing and training simultaneously (Kylli, 2005).

Through the 1990s, the educational uses of Internet were limited to e-mail and discussion forums. The only problem at that time of the wide use of these applications was the low bandwidth and not minimum access among users. The largest growth of Internet uses occurred when World Wide Web started. World Wide Web granted large amounts of data i.e. graphics and texts to be designed, sent, and exchanged (Kylli, 2005).

After 2005, a large number of new web tools started to be used widely in different aspects, specially in educational practices. The development of web applications and tools at that time is referred to as Web 2.0 tools which are characterized by being interactive models, also called the push model (Kylli, 2005).

Consequently, web applications were largely applied into learning and education practices. This era represented the beginning of what is called *off-the-self* learning environments. These environments combine course materials, exams, homeworks, e-mails, and discussion forums, into one online environment, called a Learning Management System (LMS). LMSs are classified into two categories, open course and commercial LMSs (Valiathan, 2002). The next section 2.6 provides a detailed explanation of the LMS.

Blended learning is another term marks the development of learning and LMSs. A blended learning is a concept that describe learning approach that combine both classroom learning activities with online learning applications and practices (Valiathan, 2002). Teacher's role within this approach is using computer programs in course plan, for example using PowerPoint slides for distributing learning materials, or online forums to distribute class news (Bates, 2010).

Learning approach throughout LMS relies largely on the teacher as being the main organizer of teaching plan, arranging materials, and choosing the online tools to be used within the course. Moreover, this model is centralized around teacher since he/she is responsible for setting the course plan. Teachers chose online activities either e-mails or discussion forums, learning material format, and the required readings. This learning approach is called e-learning 1.0 (Downes, 2005).

On the other hand, another form of courses are the ones which are taking place completely online. This is referred to in figure 2.1 as distance learning. The newest approach of learning and courses is what is called Massively Open Online Course (MOOC). It is a learning model which takes place for students who are virtually registered in totally online courses. Anyone who are interested in these courses can register freely; no matter if he/she is a student at university, high school,

or even retired. The only requirement is having an Internet access (Initiatives, 2011).

MOOC is characterized as a cloud computing environment in such a way it aggregates different services and applications to create and structure the course. MOOC can be defined as an open web-based learning environment composed of different cloud based services in which learning material are created and distributed among all participants (Kop & Carroll, 2011).

Moreover, MOOC is seen as a very good example of utilizing the newest learning approaches and structures. Learners within this course type are responsible for designing the course structure based on his/her learning preferences. Each learner within this course has his/her own learning environment, in other words, he builds his own a personal learning environment (PLE), or a personal learning network (PLN) by choosing the preferable learning tools and services (Mott, 2010).

The newest learning concepts, PLE and PLN, are considered as the newest replacement of the traditional LMS environments. Comparing LMS to PLE, LMS's existence is linked to learning institutions and it is referred to as course-based environments, the reason is all learning activities are held within the offered courses (Mott, 2010). While PLE is more informal approach for learning environments where the tools used in learning are chosen based on the learners preferences. Also, they are not tightly linked to learning institutions and not related to grades and assignments deadline submissions. Section 2.6 gives a more in depth explanation regarding PLE and LMS.

Accordingly, MOOC is seen as an example of connective networked learning approach. Creating MOOC course is constructed by learners through their own personal learning environment according to their abilities and educational background. It benefits from RSS technologies which accumulate different learning materials from online activities such as blogs, discussion forums, links, and wikis. Each learner combine activities and material based on his/her own preferences and most importantly suit their objectives. Later on sharing them with other learners with similar preferences and learning styles (Siemens *et al.*, 2010).

Recently, more courses were offered online in different research fields. The first example of MOOC course is the Personal Learning Environments, Networks and Knowledge (PLENK2010). It was offered by the National Research Council of Canada. It is a free course anyone around the world can register to it with the total number of participants were 1641. The duration of the course was 10 weeks (Siemens *et al.*, 2010).

There is a large amount of information generated from learners' interactions from the different activities of MOOC. Many research directions have been introduced to study and analyze these information. For example, learning analytic tools based on Social Network Analysis (SNA) were used to analyze information extracted from PLENK2010 course. This analysis is based on the assumption that the relations

and interactions between learners within the course are represented as a network. The formed network can be later analyzed and studied. The data analysis about the learners and their interactions can provide useful information about learners (Kop & Carroll, 2011).

The next section presents a detailed explanation of e-learning 2.0 and its development.

2.2 The Evolution of E-learning 2.0

The introduction and large growth of *Web 2.0* have a huge impact on the social life style of individuals and as well as in education. The use of Web 2.0 technologies and tools in education is referred to as *e-learning 2.0*. E-learning 2.0 is a concept that define a group to technologies and online activities that are used into learning practices and applications. It is seen as a continuity of both e-learning and technology enhanced learning.

The significant growth in Internet use and being easily accessible have an impact on education. The first effect of Internet on education was done through digitizing learning material i.e. texts, pictures, and videos. This is marked as the era of Web 1.0. Learning material were uploaded on the course homepages and shared with the students. The participation was restricted only to uploading and downloading content and browsing among them, however, with very minimum interaction. The main characteristic of Web 1.0 is being content-based web pages (Bessenyei, 2008).

In the recent years the increased advancement in Web 2.0 technologies have boasted learning applications and communications. The use of Web 2.0 applications in learning have generated new learning approach where learning become continuous process and not restricted to course timetable and deadlines. Also, this has defined informal methods for learning, focusing on communication between peer learners (Siemens & Downes, 2008; Dron & Anderson, 2009). Although Web 2.0 applications and social learning have been very recently applied into learning approaches, it gained a large success and have changed how learning is taking place (Gulati, 2003; Conole *et al.*, 2008).

Table 2.1: Differences between Web 1.0 and Web 2.0, adapted from (O'Reilly, 2005)

Web 1.0	**Web 2.0**
uploading	interactive
websites	blogs
content management	wikis

Table 2.1 summarizes the main differences between Web 1.0 and Web 2.0. The term *Web 2.0 technologies* describes the technologies that emerged from Web 2.0 around 2004. It focuses on interactions, participation, and knowledge sharing. Web 2.0 tools are blogs, wikis, social networks websites, photo sharing websites, and online bookmarks websites (O'Reilly, 2005).

As seen from the table, the main difference between the Web 1.0 and Web 2.0 is the instant and direct interactions between like-minded participants. In web 1.0 sites they were based on hyperlinks and accessing online data (O'Reilly, 2005).

On the other hand, Web 2.0 tools are characterized by the ability to grant participants more authorities; not only accessing online content but also sharing ideas and thoughts through an open environments (O'Reilly, 2005). In literature Web 2.0 involves these subsequent features: (1) sociable and interactive web is the main feature of Web 2.0, and (2) new tools and services used in everyday aspects of users' live (O'Reilly, 2005).

Noticeably, Web 2.0 tools have demonstrated its high usage in both day-to-day applications either formal and informal aspects. The integration of Web 2.0 services within e-learning is denoted by *e-learning 2.0* (Downes, 2005). Table 2.2 shows the main differences between e-learning 1.0 and e-learning 2.0.

The services and tools used within LMS are marked by the use of Web 1.0 applications. Most universities or learning institutions, even schools, use LMS. LMS used interchangeably with Virtual Learning Environment (VLE) or Content Management System (CMS) manage learning process in every single aspects. LMS, or CMS and VLE, share common features such as managing courses (quizzes, learning materials, etc.) for large students numbers. There are many well known examples of LMS that range from open source systems, like Moodle and Sakai, to property systems, like, Blackboard/WebCT (Dagger *et al.*, 2007).

Table 2.2: Differences between e-Learning 1.0 and e-Learning 2.0, adapted from (O'Reilly, 2005)

E-Learning 1.0	**E-Learning 2.0**
LMS	Web 2.0 applications
Managed by teacher	Managed by learner
Controlled by teacher	Teacher only observe
Communication among course members	Communication among community

Recent years have witnessed an increased importance of Web 2.0 in education. This can be seen through the use of Web 2.0 services in PLE. The use of Web 2.0 in higher education form the beginning of he new ICT pedagogy, referred to as *pedagogy 2.0* (Dron, 2006).

2.2.1 Technological Aspects of Web 2.0

Table 2.3: Uses of Web 2.0 within learning practices, adapted from (Anderson, 2007)

Web 2.0 Service	Educational Uses
Blogging	• providing instant feedback for students also among students between themselves. • using comments can motivate students to help each other in their writing as well as their assignments.
Microblogging	• creating courses community, benefiting from online courses discussions, and creating Personal Learning Network.
Wikis	• producing a course FAQ as a place for discussions and combing new resources. • using wikis for managing students' projects and organizing materials from groups of students
Video Sharing	• producing specialized videos with students and benefiting from video sharing sites in finding videos for special purposes.
Photo/Slides Sharing	• grouping of social media applications, managing community, and personal learning environments.
Syndication of content through RSS	• updating information in courses. • RSS feeds can possibly take over email lists and groups

Some of Web 2.0 services and applications are already being used in education. They are built based on open source standards. The very well known examples are blogs, wikis, multimedia sharing services, content syndication, podcasting and content tagging services. The very well known examples are blogs, wikis, multimedia sharing services, content syndication, podcasting and content tagging services (Anderson, 2007). Table 2.3 shows some examples of how Web 2.0 services are used within learning practices.

2.2.2 Social Aspects

E-learning 1.0 main purpose is distributing and administrating learning content to students. Whereas e-learning 2.0 main purpose is social aspect of learning. This extends the limited social tools within e-learning 1.0 which are restricted to emails, messages, and discussion forums (Downes, 2005).

On the other hand, e-learning 2.0 focuses on the social interactions. It is based on providing a platform for learners who share similar research interest to collaborate. There are endless number of example, one is social networks like Facebook[1], which focuses on connecting users (Downes, 2005).

Other example of Web 2.0 applications is web sharing applications (Ebner, 2007). Online example is social bookmarking tools like del.icio.us[2], which enable users to distribute and share bookmarks of similar interests. Also, websites for sharing multimedia content; videos like Youtube[3] and photos like Flicker[4].

2.3 Main Ideas of Web 2.0

The main ideas behind Web 2.0 is summarized by six items in the study by Anderson (Anderson, 2007). They are:

(1) "user generated content,
(2) the wisdom of the crowds,
(3) the long tail in learning
(4) architecture of participation,
(5) network effects, and
(6) openness."

Though, the important ones are: the wisdom of the crowds, the long tail in learning, and network effect and Openness. More detailed descriptions are as follows:

The wisdom of the crowds

This concept explains the principle of an online item importance. It states that an item's importance or popularity of an increase by the larger number of users accessing it. The older static websites do not depend in their success on the largest number of users accessing them. These websites offer the same material since they are static websites. On the other hand, Web 2.0 sites take advantage from users feedback and frequent access to the website. These websites use the previously mentioned information to modify and update their content and the presented information. For example: Amazon[5] and Wikipedia[6] (Anderson, 2007).

[1] www.facebook.com
[2] http://del.icio.us
[3] www.youtube.com
[4] www.flicker.com
[5] www.amazon.com
[6] www.wikipedia.org

The long tail in learning

It defines the concept that the recurring occurrences take place by a power law distribution. A power law distribution plot with small number of occurrences at the end. Relating this concept to learning applications can be observed in two dimensions: first, generating and building online material, second the simple method of publishing these content online. Consequently, this enable the easy interaction through sending and receiving feedback and comments (Redecker, 2009).

Network effect and Openness

The large evolution of the web and its technologies have been accompanied with development of the copy rights of digital learning content. Hence, web technologies have been developed using open technological standards, like open source software. As an example of open source browser is Firefox which offers free updates through added plug-ins (Anderson, 2007).

2.4 Examples of Web 2.0

As stated in the previous section 2.2.1, the majority of Web 2.0 applications are focused on user's role being interactive throughout content generation, distribution, and collaboration and cooperative with other peers. Basically, collaboration and social interactions form the basis of Web 2.0 applications. *Web 2.0* is defined as:

> Web 2.0 is the network as platform, spanning all connected devices; Web 2.0 applications are those that make the most of the intrinsic advantages of that platform: delivering software as a continually-updated service that gets better the more people use it, consuming and remixing data from multiple sources, including individual users, while providing their own data and services in a form that allows remixing by others, creating network effects through an architecture of participation, and going beyond the page metaphor of Web 1.0 to deliver rich user experiences (O'Reilly, 2007).

One example that assembles almost all Web 2.0 applications into one single example within a learning context is the Personal Learning Environment (PLE). The PLE term is the newest arrival in online learning environments. Section 2.6.2 presents a more detailed description of the PLE concept and its examples.

The integration of Web 2.0 into learning practices introduced new ways of learning. For example, *informal learning*, *lifelong learning*, *personalized learning*, and *network learning*. They all share the main idea of integrating Web 2.0 and social media into

learning contexts, but each term has its different and common points. The next subsections give a deeper more detailed explanation of these terms.

2.4.1 Informal Learning

Learning approaches and methods which take place under the regulations of learning institutions are marked as traditional approach. These traditional approaches are referred to as *formal learning*. It is called formal approach because learning takes place based on the structure of learning institution, i.e. schools or universities. The teacher has the control over creating courses and initiating teaching approach. Formal learning approach has the course as the central building block along within the learning environment. Teachers have the main role is to design courses and is responsible about grading students performance (Chatti *et al.*, 2007).

Nevertheless, an important argument in learning is the nature of *informal learning* or *non-formal learning*. Klamma states (Klamma *et al.*, 2007): "once you step beyond traditional institutional boundaries you can find learning that is driven by and for you, the learner". Informal learning is not restricted by a place, time constraints, or exams and grading. This type of learning is characterized by being available everywhere and anytime, and most importantly the plethora of information and knowledge that can be accessed.

Cross (Cross, 2006) claims the fact that students learn the best outside class boundaries and within friends circles. Formal learning methods form only 10% to 20% of the methods used for learning. Learners learns more through experiments and making mistakes not only from reading books and attending classes. Informal learning concept point out the idea that learning is not restricted to learning institutions and take advantage of informal sources. For example, using games, experiments, online tests, discussion forums, and social network sites, all of these sources allow learners to work jointly on similar problems and share ideas and thoughts (Chatti *et al.*, 2007).

2.4.2 Lifelong Learning

Brown and Adler (Brown & Adler, 2008) declared the idea that people do not rely on the knowledge they gained from universities many years ago once they join new job. On the contrary, they use the experiences and practices they acquired throughout their career. Learning does not stop after finishing college nonetheless it is a never ending process and the acquired knowledge increased over time. This idea is more highlighted in these days with the explosion of information and online applications where anyone can continue their learning.

Another recently emerged concept is *lifelong learning* (LLL). It refers to the continuous availability to learning possibilities for anyone who wants to learn without any

constraints regarding grades (Aspin & Chapman, 2000). Basically, learning is an activity that not limited to formal approaches; though it is a constant activity that goes during lifetime. In Web 2.0 era, lifelong learning is getting more trendy and challenging. Most of the traditional educational systems did not alter their learning models to go with the new developments in online methods. There is an emerging need for new learning models that derive the need for lifelong learning (Klamma *et al.*, 2007).

2.4.3 Personalized Learning

One main feature in learning is the very personal experience of every individual and the way he/she prefer to learn. Every learner has a unique style that is a reflection of his/her characteristics and their own knowledge. Personalized learning is another new term that describe the idea that each one choose learning model that is very appropriate.

Attwell (Attwell, 2007b) suggested empowering learners to choose their own learning styles along with supervising them according to their own capabilities. It also emphasizes on making the student the center of his/her own learning process. Jafari (Jafari *et al.*, 2006) further that learners need to have a system that is centered on what they want and need. This requires this system to be adaptive and responsive.

Personalized learning sets a learning model that not tied to the curriculum and one form that fits all the learners regardless their needs and preferences. This new model focuses on students and they are the center of the learning process. It is also referred to as a learner-centered approach.

2.4.4 Network Learning

Learning is an process based on group activity and connections among group members. These connections form a network and accordingly learning lies in the network connections (Downes, 2006; Siemens, 2006). Network learning is a concept driven by social medial and Web 2.0 applications and has raised highly recently.

Nowadays, web users have the privilege of accessing information from many different sources over the Internet. These sources are characterized as being available anytime without any restrictions. The new era of Web 2.0 is characterized by the highly social interactions between individuals regardless of their physical location. This have a great reflection on education in such a way that learners have much closer and immediate contact with their peers which make them share ideas, establish new connections, and share common problems (Downes, 2006).

A new generation of learners evolved due to the changes of learner's perception of Web 2.0 applications. The new generation in education are not merely seen as consumers but also as active producers of knowledge. Each learner have an essential role in choosing the tools and methods in which he/she are comfortable with. The numerous options of learning material sources through the web along with its social collaborative, participative nature result in a new phase in learning mainly described by networking and sharing (Downes, 2006; Siemens, 2006).

2.4.5 Social Learning

The integration of social networks and education has been referred to as social learning in which its core is exchanging, transferring, information among like-minded people. This type of learning is the contradictory of the traditional institutional educational systems. The application of social learning is centered around not having a one model applied to the whole course but a flexible one that is accumulate different online tools and applications. Another main feature of the social learning is the importance of connections, in other words, the learning peers whom share the same research interest. Many new learning platforms offer the opportunity to connect to other peers within the defined learning network (Vassileva, 2008).

The discussed theories of *constructivism*, *situated learning*, and *social learning theories* have a large impact on Communities of Practices (CoP) approach. CoP concept is one of the earliest concepts which set the foundation for knowledge sharing (Li *et al.*, 2009). CoP defines a system in which individuals who share similar research interest grouped together. It focuses the fact that individuals benefit more through social and informal relations that are formed among communicating with their peers (Wenger, 1998; Rogers, 2000). Based on the social learning theory defined by Wenger (Wenger, 1998), communities or groups of similar interests are defined from the participation and interaction between them. The larger the community is, the better outcomes are gained by students in terms of finding answers to common problems.

2.5 Critics and Discussions

The previous sections presented an in-depth review for some basic principles and examples of Web 2.0. Researchers believe that there will be a very huge boost in the uses of Web 2.0 in almost every aspect of daily life, more specifically, in learning context. E-learning applications take advantage of the collaborative nature of Web 2.0 services i.e. collaborative content creation. Most importantly, the highly use of the concept of learning network along with introduction of many examples them. Learning networks and e-learning 2.0 have attracted large attention lately and become and important research topic (Neal, 2007).

Original Web 2.0 services were not solely designed for educational uses, but their applications and the ease of use, have drawn a large attention in learning environments. The main features of Web 2.0 are the availability of large amounts of information, also in different formats i.e. videos, photos, audios, interactive tests, instant feedback, and the ease of use. All of these features have a great impact on provoking new strategies in teaching such as encouraging interactive learning, empowering students in their learning process, and focusing on social learning (Downes, 2006).

In other words, applying Web 2.0 services into teaching practices and methods have many advantages for both teachers and students. These advantages can be seen from Web 2.0 services such as blogs and wikis which encourages students participation and engagement within the course. Also, benefiting from social networks in many aspects for example creating a profile for each student and allow them to participate in learning networks of specific type. All these services encourage students to participate more and being more active either to contribute in wikis, engaging in conversations in discussion forums as an example (Grosseck, 2009).

Web 2.0 services have expanded the class boundaries into a virtual environment. Referring to the course as a virtual environment enable larger number of students to enroll and encourage them to seek knowledge outside the small course boundaries. This definition of a course enlarge its idea to a whole virtual environment. This motivate students highly to engage with larger groups and the benefit of exchanging ideas with larger communities from different backgrounds. Moreover, using Web 2.0 services have no time constraints and is accessible all day long. This allow students to learn, find solutions, and interact with their peers within their virtual environment around the clock (Grosseck, 2009).

Many studies showed a strong supporter of applying Web 2.0 into the learning context. According to (Grosseck, 2009), Web 2.0 importance stem from its collaborative nature which motivate learners to produce and exchange learning content and accordingly sharing them with each other. Moreover, this study stated course syllabus must be designed based on the new learning theories and standards of Web 2.0 services. This defines the new educational era.

Another study by Meyer (Meyer, 2010) examined how applying Web 2.0 services among doctoral students research work. It studied the use of wikis, blogs, and online discussion forums among writing and discussing and jointly writing a research papers. It is found that students were willing to use Web 2.0 services and referred to its importance in sharing ideas. Additionally, students revealed that using the online spaces are easier and more practical than the traditional learning environments. However, small group have some concerns regarding using the new learning approach.

Despite the fact that Web 2.0 services have many advantages, there still are doubts on applying them widely on education. These doubts must be further investigated in order to guarantee its fully use into education sector. One essential doubt that

hinder the fully usage of Web 2.0 services is the privacy issues. The idea of sharing personal information of participant i.e. teachers and students, among the web and each other (Grosseck, 2009).

2.6 Learning Management System and Personal Learning Environment

A Learning Management System (LMS) – also called Course Management System (CMS) or Virtual Learning Environment (VLE) – are systems that are controlling, supporting, and organizing courses along with the communication between students and teachers among them. In general, LMS can be defined as a software or webbased application used to organize, implement, and apply online courses. Moreover, to organize and distribute learning material and process among student (Alias & Zainuddin, 2005).

Specifically, LMSs grant the availability of learning material among course members. LMSs can support two types of courses, the online course i.e. distance learning, where learning done totally online. The second type is partly online courses where the learning take place face-to-face but the course organization done online through LMS. Additionally, LMS enables teachers to organize assignments, tests, and supporting learning tools such as discussion forums and wikis. Finally, the main feature of LMS is allowing collaborative learning and providing an environment in which students work together and provide their contribution to the teaching process (Romero *et al.*, 2008).

LMS is classified into two main categories, either open source or commercial systems. Some examples of commercial systems are Blackboard/WebCT[7], TopClass[8], etc. Some examples of open source systems include Moodle[9], Ilias[10], Claroline[11] (Romero *et al.*, 2008).

Nowadays, one of the largely used systems is Moodle (Modular Object Oriented Developmental Learning Environment). Moodle is an open source LMS which offers an online environment for creating online courses and organizing collaborative group among registered students (William, 2006).

There are many different types of LMS that are used in these days. All have one objective which is providing online collaborative learning environment. They have some common aspects, such as (Oneto *et al.*, 2009):

[7] http://www.blackboard.com/
[8] http://www.wbtsystems.com/
[9] www.moodle.org
[10] www.ilias.de
[11] www.claroline.net

- organizing courses and participants among their roles i.e. teachers, students, tutors,
- managing course timetable with deadlines and tests dates,
- composing the course plan through defining the online tools and activities to be used, and
- exchanging messages among registered students and their teachers and tutors.

From a technical point of view, LMSs are developed using tools like Java based frameworks, Microsoft .NET, PHP. All these systems are combined with database systems to manage large amounts participants information. However, the commercial LMS have software license and limited access (William, 2006).

The Technical University of Berlin (TU Berlin) uses Moodle as its LMS and is called Information System for Instructors and Students (*ISIS*)[12]. Section 5 provides in depth look at ISIS, as it is used as a case study in this thesis. The following section provides a more detailed look at Moodle.

Some other terminologies are used in the context of using computer technologies within learning, they are: Computer Supported Collaborative Learning (CSCL) and Computer Supported Collaborative Window (CSCW). These two concepts define how computers and online technologies can be used to enhance communication and collaboration among students regardless of sharing one physical location. Also, they emphasize on the importance of the social describe how computers can be used to facilitate social relations between students for better learning outcomes. CSCL and CSCW have one objective is to enhance group interactions and team work among students Haythornthwaite (1999); Cho *et al.* (2007).

2.6.1 Moodle

Moodle stands for *Modular Object Oriented Dynamic Learning Environment*. It is an open source virtual learning environment designed based on a learning style called *Social Constructionist*. This learning style states that for better learning outcomes, students have an active role in their learning process. This means students need to contribute and define the learning material and engage in discussions with their peers about the learning process. Nevertheless, Moodle can be used as a content management system regardless of its constructivist approach (William, 2006).

Moodle's main objective is establishing an interactive online learning environments as an addition to improve the in class learning. Moodle is considered one of the highly used systems; its users have increased largely from its first release in 2000. The statistics from October 2012 showed the number of registered verified sites have exceeded 70.793 over 224 countries that serve 6.7 million users (Moodle, 2013).

[12] www.isis.tu-berlin.de

Moreover, Moodle system allow teachers to create resources and activities to support learning process. Moodle is a course based learning environment, this implies that all the resources and activities are accessed within the defined course. In order that students participate in these activities or access the resources, they have to be registered within the course. For example, the activities are: assignments, chats, questions, discussion forums, and glossaries (William, 2006).

Other activities include quizzes, surveys, wikis. Another resource among Moodle is Sharable Content Object Reference Model (SCORM). It is a resource that allow Moodle to create and aggregate a learning content that a content from different types that can be presented later on different learning environments (William, 2006).

Moreover, Moodle is implemented based on LAMP open source framework. It is composed of Linux for the operating system, Apache for the Web server, MySQL for managing the database, and PHP for programming the interface (William, 2006).

2.6.2 Personal Learning Environments

Nowadays, there is a new trend in online learning called personalized learning. It is an approach which grant learners more responsibilities by allowing them to choose the learning style that suits their abilities and preferences. These personalized learning environments can be seen as an aggregator of many services that students are comfortable using them. This new concept is a result of the large growth in Web 2.0 services and applications (Attwell, 2007a; Downes, 2005).

Personal Learning Environment (PLE) is a new term which represent the new approach of personalized learning. PLE differs from the traditional learning approaches in which all the activities are taken place within course and among registered students only. The traditional learning approaches have one common learning plan regardless of students' preferences or styles (Attwell, 2007a).

PLE is rather a new term, first introduced in 2004 (van Harmelen, 2006). Van Harmelen described PLE as a new learning environment that is composed of different educational resources, services, and applications that are organized by students themselves. This allows students to create their own personal profiles that reflect their learning interests and approaches (van Harmelen, 2006).

As it seen from the definition, the advantage of PLEs resides in being centralized around learners themselves. This differs fundamentally from the LMS solutions. It is a way of granting students larger roles in designing their own learning approach through choosing, aggregating, distributing different learning resources that construct their personal environment (Tery Martindale, 2010; Attwell, 2007a).

PLE can be defined, from a technical point of view, as a software that aggregate many different services. It is not a single tool, on the contrary is a combination of Web 2.0 services that form a single PLE (van Harmelen, 2006)

Downes (Downes, 2005) extends PLE concept and points out the differences between the traditional i.e. institutional learning systems and PLE. First, the learning applications forms a small personalized learning environments controlled by the learners themselves. The content and the type of the services are chosen based on the learners' preferences rather than being defined by learning institution. PLE are indeed environments and not only systems because they are composed of collection of collaborative operated services.

Figure 2.2 is one of the earliest diagrams describing the concept behind PLEs and their relation to LMSs. It is called *Future LMS*, which was introduced by Scott (Wilson, 2005). This figure focuses on the idea that the future LMS will not only be a repository of learning materials but as an aggregator of the services and applications that are comfortable using them.

This model illustrates how the *future VLE* will act as a personal website that combine tools and services for learning. The figure 2.2 display that the future VLE is a combination of different applications, like 43Things[13], LiveJournal[14], or Delicious[15]. This new definition of VLE enlarge the strict definition of the old one. Most importantly, the future VLE does not use the services offered by the institutions but it involves services from learners' choice, i.e. online services based on their preferences. Thus, the new environment benefits from online services APIs (for example, Google Docs[16]) that generates interactive interfaces .

The future VLE model enable learners to generate and register to learning material for their studies based on their preferences. This is done through utilizing online service such as Rich Site Summary (RSS), Friends-of-a-friend (FOAF), and my other web services. The idea is not only accessing materials but also distributing and exchanging them with other peers to enhance the social interactions between learners. This creates communities of learners with similar research interests and with focused approaches and can enhance their outcomes 2.2.

Moreover, the concept of future VLE has been developed into PLE. PLE has extended the concept of the classroom into a whole virtual class with open services and learning resources. PLE main focus is decentralization of the learning process and openness of resources and services. It allow learners to benefit from the endless options from Web 2.0 services. As Downes (Downes, 2006) states: "the heart of the concept of the PLE is a tool that allows any learner to engage in a distributed environment consisting of a network of people, services, and resources".

[13] www.43things.com
[14] www.livejournal.com
[15] https://delicious.com/
[16] docs.google.com

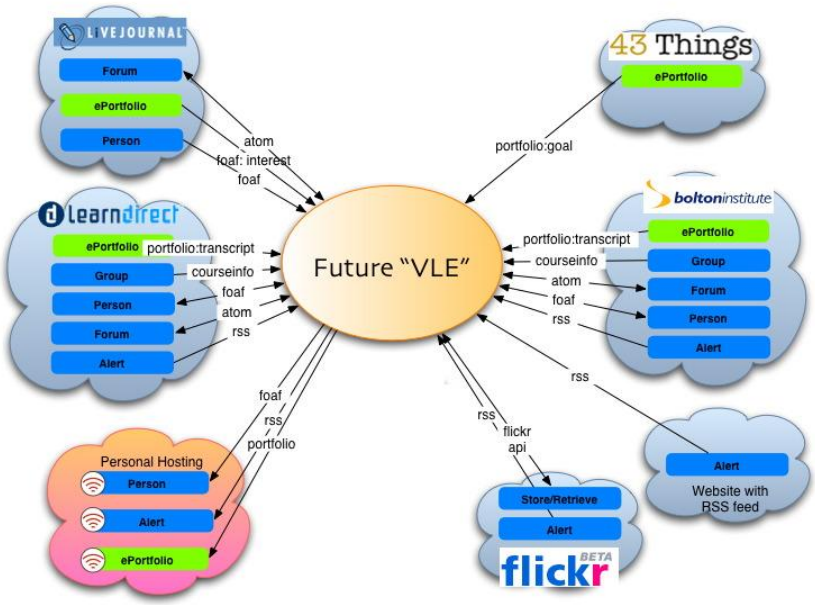


Figure 2.2: VLE of the future (Wilson, 2005)

Another difference between LMS and PLE resides in the fact that LMS are based on firm separation between the roles of teachers and students without any overlapping between them. This can be seen on the direction of information distribution i.e. services and resources used, from teacher into all students in the course. PLEs however have the opposite direction of information flow, guided by learners' choices of services for their learning styles, rather than being imposed by teacher's plans (Wilson *et al.*, 2006).

Many researchers have investigated and defined PLE concept. For example, Jafari (Jafari *et al.*, 2006) depicts the need for new learning environments as: the traditional LMS/CMS, in their conception as management system, ends the interactions between learners and teachers by the end of the semester. The social relation and information exchange is finished outside the universities boundaries. A PLE, however, combine all the new learning approaches formal, informal, and lifelong learning within one environment centered around learners.

Furthermore, Attwell (Attwell, 2007a) explained PLE as an online learning platform that associate all the newly introduced learning concepts and approaches informal learning (through social media), practical learning, learning in free time (without institutional restrictions), and learning methods based on the individual styles and preferences.

A list of other examples of implemented PLEs or aggregated Web 2.0 services for learning environments is following below.

Personalized Startpages

The number of personalized startpages has increased exponentially through the introduction of Web 2.0 applications. They offer an online platform where users can generate and aggregate information and knowledge regardless of their physical location and anytime throughout the day. These startpages are one of the tools which can be customized to build a PLE. Also, these pages can help learners to gain a centralized part in defining the learning methods, and accordingly allow them to choose the online technologies most appropriate to them (Ivanova, 2009).

Integrating Web 2.0 services into learning approaches have revolutionized the method of aggregating data from various online services. Most specifically, mashup technologies, from Web 2.0, have allowed PLE environments to be built based on mashup technologies which allow services aggregation. is examined in order to see what kind of components can be used in PLE implementation. Startpages create an individualized learning environment which serve as a learning model for each single learner. Examples of these startpages include Netvibes[17] and iGoogle[18].

PLEF

The Personal Learning Environment Framework (PLEF) is a platform for creating personal learning environments by individuals that expand the traditional learning approaches. It is built based on mashup system which allow learners to create their own environments by aggregating online services from different sources. It enable learners to use customized services that are all combined in one platform. PLEF offer learners to use web services technologies like REST and AJAX to aggregate their selected services (Chatti, 2010).

Furthermore, PLEF provide learners with some services in order to build their individual PLEs, like: aggregation, tagging, sharing services (e.g. feeds, widgets). Another feature of PLEF is that it has a navigation service that allow learners to drag and choose the services between to create the different pages of their PLE. PLEF also allow search option within the environment either by full-text search or tag one (Chatti, 2010).

PLEX

PLEX[19] is another example of PLE developed at the University of Bolton. It is developed within the Personal Learning Environment Reference Model project[20],

[17] www.netvibes.com/en
[18] www.google.com/ig
[19] http://www.reload.ac.uk/plex/
[20] http://wiki.cetis.ac.uk/Ple

which is funded by the Joint Information Systems Committee (JISC)[21]with the learning Frameworks and Tools Programmed in 2005. From a technical point of view, it is built from these application: RSS subscription, FOAF, and a organizer for favorite web links.

Elgg

Elgg is another example of PLE system that has an increased importance lately. It is an open source online system offer an online social platform for educational purposes. It is built upon the following web applications, like webblogging, e-portfolios, and social networking. Elgg allow learners to form their learning communities in which they can exchange and distribute information among their peers (Campbell *et al.*, 2005).

The defined communities can improve and encourage students even outside classroom learning. These communities can defined for specific learning purpose like learning a foreign language. Within these communities, learners can contact their peers, share learning materials i.e. text, photos, videos and also exchange ideas within discussion forums (Campbell *et al.*, 2005).

Elgg offer each user their own weblog, file repository, their own profile, and RSS service. Learners can be connected with others who share their research interest by tag collection of keywords. These tags can define learning networks within Elgg (Campbell *et al.*, 2005).

Blogs

A blog, or online journal, is an online personal website which involve users to apply their own material and information. Blogs are considered personal websites where learners can customize its use according to their own defined objective. Individuals can upload videos, graphics, and even exchange comments with others who subscribed within the blog. The uploaded input from individuals are presented from the newest into the oldest ((ELI), 2005b).

Many websites offer the service of creating online blogs. One example is Blogger.com[22], which gained a lot of popularity due to its simplicity. Blogs have been applied in many approaches, either personal, business, and educational uses. The simplicity is seen in the ease of sharing information through uploading individual's thoughts and discuss ideas through commenting system.

As stated earlier, the main characteristic of blogs is being easy accessible. The challenge is integrating it in learning plan, and whether it is useful to be applied. It

[21]http://www.jisc.ac.uk/
[22]www.blogger.com

is suggested that it is useful in critical thinking applications. Also, the instructor can define a blog for the course purposes where the course information and news can be distributed among participants (Williams & Jacobs, 2004).

2.7 Learning Theories

This section studies the concepts of learning theories and their relation to the traditional and new learning environments. A learning theory illustrate how students behave within the learning environments and the best method to learn. In literature, there are different types of learning theories: *behaviorism*, *cognitivism*, *constructivism*, as well as the newest one *connectivism*.

Almost any learning environments, and especially LMS, are developed based on learning theories. In the following subsections more information is provided.

Behaviorism

Behaviorism is one of the earliest defined learning theories. It is based on monitoring how learner's behavior change with obtaining new knowledge. This learning theory is based on providing a motivating leaning factor for learners and later study their reaction to this entered motivation (Jung, 2001).

Learners within this learning theory are responsive to the introduced conditions within the learning environment. Therefore, this type of theory largely focus on the circumstances that create the learning environment. An example of a learning application that apply behaviorism is Computer-Assisted Instruction (CAI) (Jung, 2001).

In order to design a learning material based on this learning theory, the following factors should be considered: (1) the learning material must be arranged in a consecutive approach in order to reach the learning goal at the end, (2) the evaluation process must be took in all gradually among the steps of the learning process not only at the end. This is important for giving comments and recommendations for better behavior and results (Anderson & Elloumi, 2004).

Cognitivism

Cognitivism learning theory emphasizes on the structural knowledge gaining. It studied how knowledge is how information is acquired, arranged, and processed by learners. Learners are described by having an essential active role within his/her learning process (Jung, 2001).

Cognitive theory focuses on the encouragement to learn new idea and how they are reflected on learners themselves. The cognitive theorist state that learning process depends on these assumptions: learners' abilities and background, the efforts done for gaining new knowledge, and the design of the learning process used by them (Craik & Lockhart, 1972).

The learning material based on this theory take into account the intellectual level of learners. This motivate students to interact largely and encourage them to study. The design of the material include mandatory quizzes to refresh their previous knowledge encourage. The learning material must also take into account the diverse learning styles of individuals, this is done by presenting different types, i.e. text, audio, video, and interactive quizzes. All of that help learners regarding of their background to acquire new knowledge (Anderson & Elloumi, 2004).

Constructivism

Constructivism is built upon the assumption that each learner must design and compose their learning process based on their individual background knowledge accumulated throughout their lifetime. Behaviorism and cognitivism theories are depend on teachers to create learning plans and design learning material. On the contrary, constructivist theory state that designing learning is centralized around learners themselves (Jung, 2001).

Learning according to constructive theory depend mainly on learners themselves. learners need to be initiative and taking larger part in their learning process. Empowering learners motivate them to participate more. Teachers on the contrary have an observant and guiding part in the learning process. Moreover, learners are encouraged to find new ideas and solutions. Therefore, learning has changed from being centralized and initiated from teachers into learners themselves (Lee & McLoughlin, 2007).

Anderson and Elloumi (Anderson & Elloumi, 2004) stated that learning is an interactive process. This theory is the basis of the idea of personalized learning. This learning theory emphasizes on the concept of participative part of learners. For example, presenting the learning material in an interactive way, like presenting learning material by interactive question answer method.

Social Constructivism

Social Constructivism theory concentrate on the collaborative and cooperative nature of learning which combines teachers, learners, and tutors. This theory gives a large importance of the social part of learning where all learning peers communicate in informal way to exchange ideas. The idea of social learning introduces the concept of group learning where the teacher's part is advisory. Moreover, this theory takes

advantage of the technological development of in forming social learning groups where learning mainly takes place through exchanging ideas, information, joint problem solutions, and background knowledge (Dougiamas, 2007).

Moodle learning environment is developed based on the idea of social constructivism. Moodle's design emphasizes the social interactive aspect of learning where learners participate in composing and creating learning materials. The participative aspect means learners acquire new knowledge from interaction based on their background knowledge and through interaction within their like-minded peers. This learning theory is considered as a development of constructivism by highlighting the importance of social group working (Dougiamas, 2007).

Connectivism

Connectivism states that learning reside in networks. These networks are formed from the social interactions between the learning participants. This theory emphasizes on the importance of the finding new connections. Knowledge acquiring resides in the network connections. The learning process begins with establishing and finding new connections (Siemens, 2005a).

Connectivism is a very new learning theory introduced by Siemens (Downes, 2006; Siemens, 2005b). It is the new theory of the Web 2.0 era that is characterized by the ease of use plenty of available information. Based on connectivism learning is a process of building connections that enable learners to acquire knowledge and learn more (Siemens, 2005b).

Connectivism is the theory of learning through establishing connections. It is based on the following combined assumptions, they are (Siemens, 2006):

- Learning and knowledge building process require diverse points of views.
- Learning mainly occur on forming networks which combine different participants from different backgrounds.
- New knowledge can be acquired within network connections resides in networks and their connections.
- Obtaining the newest updated knowledge format is the goal of connectivist learning theory.

According to Siemens (Siemens, 2005a) the importance of network connections can be seen in keeping participants closer to their peers. This means always obtaining updated information and news. The updated information can be found in the connections and not always in the immediate circle of connections.

Furthermore, Siemens find out that defining learning networks is a method for aggregating up-to-date information from different sources especially with the large amounts of information available on the web. Siemens also stated: "learning networks

can be perceived as structures that we create in order to stay current and continually acquire experience, create, and connect new knowledge (external)" (Siemens, 2006).

Connectivism is considered the newest theory in social learning, and based on learning networks. It also mark the beginning of newest learning approaches such as *social learning*, *informal learning*, and *lifelong learning*. These new approaches highlight the importance of building communities in acquiring new knowledge. Basically, it is a reflection of the development and increased use of Web 2.0 and its application in e-learning 2.0 (Siemens, 2004).

Additionally, connectivism forms the foundation of the new learning environments such PLE, Personal Learning Networks (PLN), and networked learning. All these new concepts state the significance of informal social encounters between peers, especially in exchanging information. Therefore, learning process has evolved into a continuous process that does not stop with the end of the class and stays in the network connections (Bessenyei, 2008).

Learning networks, similar to social networks, is composed of participants who share similar interests. Learning networks are a support for the formal learning approaches i.e. universities (Berlanga *et al.*, 2009).

2.8 Networked Learning Model

The previous sections studied learning environments, learning theories among with the new developments in the web. Also, it studied the introduction of e-learning 2.0 as the integration of Web 2.0 services into e-learning practices.

Figure 2.3 studies the basic components of learning theories and considers the new components of the new model of networked learning. It combines the new emerging theories of connectivism, personalized learning, and social learning into one the new networked learning model.

The networked learning model represent the new sociable learning model which is based on connecting different learning elements and participants (Steinert, 2010).

The networked learning model is designed as a result of the development of e-learning 2.0 and Web 2.0. It is built on three principles: connectivism, social network analysis, and collaborative learning (Steinert, 2010). First, based on connectivism theory, finding connections and establishing new ones are essential for the learning process. This model also emphasize the importance of connections in acquiring knowledge and learn more. Learning is an interactive process that is based on the interactions between learners and teachers (Steinert, 2010).

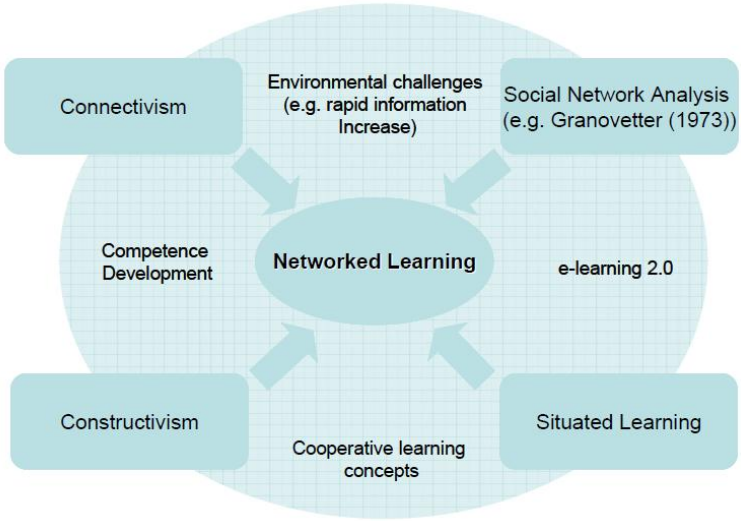


Figure 2.3: Networked learning concepts (Steinert, 2010)

Secondly, from the social network analysis perspective, a network is a set of people, or groups which are referred to as nodes or entities within the network that are connected with each other. In this case, studying the relations between the entities are more important than the entities themselves and their characteristics (Kop & Hill, 2008).

3 Chapter 3 Social Network Analysis (SNA)

The idea of studying and analyzing any phenomena by representing it as a social network has been strongly used applied in social sciences applications. Recently, it has been applied to e-learning problems, for example, as a tool for analyzing different learning environments or courses activities. The idea of studying any problem from a network perspective can not be seen as a new one since networks are seen all around. Hence, networks are all around since people are connected either with their team mates, colleagues in workplaces or schools or universities, and studying groups. Networks are the key for analyzing the problem under study.

This chapter therefore presents an overview of Social Network Analysis (SNA) along with its measures, and illustrates how important it is to use them to analyze learning environments. Moreover, it provides a brief introduction of graph theory basics, and describes the main relation between SNA and graph theory (through network patterns and visualizations). It also shows the relation between Data Mining (DM), Educational Data Mining (EDM), and Learning Analytics (LA). SNA is considered one of the applications of DM and EDM.

3.1 Introduction and Definitions

This thesis utilizes SNA to study students' learning styles, roles, and performance within online learning environments i.e. hybrid learning environments. Moreover, it explores network structure and patterns that represent the different types of relations formed from students interactions with each other and with the teachers in the learning environments.

Moreover, SNA is considered one of the applications of of learning analytics which are used to investigate the relations which took place within online courses. This gives an overview of how students participate in the online environments and their learning styles in general. SNA applications and measures grant many advantages in understanding learning relations within online environments (Haythornthwaite, 2005).

SNA was adopted in many research fields to study different types of relations and problems. Examples of these research fields are physics, life sciences, computer science, statistics, and social sciences. The foundation is to interpret and network structures and their models. Applications of SNA and their problem solving approaches include studying and analyzing social networks, the World Wide Web, and recommendation systems (Haythornthwaite, 1999).

Lately, the interest in using SNA applications has increased largely and been applied to new research fields problems. It's uses are based on extracting information and later forming the network which represent the relations between the interactive participants. SNA applications provide two main levels of analysis for studying the social relations which range from studying the whole network to studying the characteristics of individuals that composing the network (Haythornthwaite, 2005).

Choosing the level of analysis depend on the focus of the study and the types of relations needed to be understood. For example, it is essential to discover the subgroups created within the network since these subgroups or communities are composed of individuals who share common learning behavior or interest. It is also important to study the individuals' position within the network along with their direct or indirect connections. Therefore, interpreting the analyzed results can give an overview about the overall network structure, the roles each individuals occupy due to their position within the network (Haythornthwaite, 2005; Müller-prothmann, 2007).

The exponential growth of social network services is considered as a result of the huge technological evolution in Internet applications and services. This is marked with the introduction of Web 2.0 services and applications. These new applications added a new aspect of using online websites and how individuals participate in them. The most important feature is enabling students to generate online data and share them with other participants. These services are referred to as collaborative services. E-mails were previously the main communication platform, but as a result of the Web 2.0, many other platforms were introduced. For example, photo sharing websites, discussion forums, wikis, and many more (Budka & Mader, 2006; manoj Parameswaran & andrew whinston, 2007).

Additional examples of online services include: chats services, social networking services i.e. blogs, wikis, and discussion forums, all of these are applications for enhancing collaborative communications. These services along with their content and interactions generate a large volume of data and personal information about the participants. Social networks sites and services gained a lot of attention and popularity lately. They are the development of Web 2.0 platforms (Downes, 2005).

Therefore, studying and analyzing these information can provide a deep understanding of the characteristics of their online environments. This leads to a growing need for data mining applications and tools in order to provide a deeper analysis possibilities

of these online platforms and the interactions that are taking place among their participants (Downes, 2005).

A *Social Network* is a graphical representation of individuals i.e. nodes and their connections i.e. relationships and links. The social structures represents also groups or communities who share one interests or research interest. Social network theory allow studying the formed social networks under many measures like degree, betweenness, closeness, or communities characteristics. These measures facilitate understanding the overall structures of the social networks (Passmore, 2010).

Social Networking Services (SNSs) are the new hype of this era, there are endless number of examples like Facebook[1], LinkedIn[2], and many more. These services organize and understand social relations among individuals and the social relations investigate different types of relations like friendships. Moreover, social networks are not only restricted to personal social relations but also include academic relations. Day after day these social networks have involved in almost all social aspects people lives (Passmore, 2010).

Social Network Analysis (SNA) is the study of the relations between individuals among representing them as nodes of a network and the relations or connections among them as links. This is an explanation of social theory for studying any relation as long as it is represented as social network. SNA is considered a very important research direction to study social relations in many different research aspects. It allows understanding of different this social phenomena through different measures like centrality, density, cliques, etc. (Passmore, 2010; Wasserman & Faust, 1994).

Network study, according to Kocke and Yang (Knoke & Yang, 2008), represents the essential need for analyzing social relationships. It studies the social connections that are formed between individuals among any defined phenomena. Social network theory is defined based on the following criteria:

1. explaining and analyzing overall social structure are generally provide full understanding of the social phenomena other than studying the single characteristics of each connected individuals,
2. the network patterns formed from actors connections have a large impact on their roles and behaviors within their environments, and
3. the formed social relations must be studied as a changing process since social relations are not static (Knoke & Yang, 2008).

There are different types of relations and individuals that can be studied and analyzed using SNA. These relations varies from families and couples relationships into relations in workplace organizations and many others. This provide a general understanding of the characteristics of the problem under study which finding solutions for problems and improving individuals' roles (Pescosolido, 2011). Throughout

[1] www.facebook.com
[2] www.linkedin.com

this chapter, examples about different types of networks and their relations are discussed.

Social networks services and sites are not the focus of this this thesis. Nonetheless, SNA measures can be utilized to study any phenomena as long as the interactive individuals are represented as nodes within a network and the connection between them is represented as link or edge. SNA, as a research method, is developed based on the criteria that any phenomena can be seen as a social relation among the patterns that are formed from their interactions.

Social network analysts focused on exploring the social relations among humans and understand their behaviors and actions. They emphasized the importance of studying the social structure from the interactions. These network structures are essential to comprehend all the aspects of the social phenomena under study. This is described as follows by (Wetherell, 1998):

> Toward a social network approach in the past two decades, the social network approach has transformed the study of contemporary communities and kinship relations. Most broadly, social network analysis: (1) conceptualizes social structure as a network with ties connecting members and channeling resources, (2) focuses on the characteristics of ties rather than on the characteristics of the individual members, and (3) views communities as personal communities, that is, as networks of individual relations that people foster, maintain, and use in the course of their daily lives.

The large boost in applying SNA measures into many research problems have accompanied with graphical software which provide a visualization of the formed networks. Social Network Analysis Software (SNA Software) provide a quantitative and qualitative analysis approaches for the extracted networks. SNA software allows studying the extracted networks through mathematical and visual analysis approaches. The extracted networks are represented into another data format that can be later mathematically analyzed and represented graphically. The basic network characteristics can be measured from the extracted network. The extracted network are represented in an adjacency matrix which represent the nodes (or individuals) and their relations and features. Mainly, SNA software packages are classified as: packages with GUI for analysis and visualization, or ones need programming abilities (Knoke & Yang, 2008).

3.1.1 History of SNA

SNA research approach first appeared in 1930s and ever then it has evolved into an analytical, statistical, and visual approach to explore the characteristics of individuals under study and their relationships. A large and diverse research fields utilize SNA as a research methodology to understand the relationships among their individuals,

as examples of these research fields are sociology, anthropology, computer science, engineering and information technology (Borgatti *et al.*, 2009).

Lately, there has been a lot of interest in SNA and its applications. This is driven by the following factors: (1) All aspects of daily life habits and applications have been computerized i.e. everything is digitized. This generates a large amounts of online data saved in large databases which contains personal information about users and preferences. (2) The high technological advancement in both hardware and software applications facilitate the ease access of Internet all the time. (3) The beginning of open source systems and free online data which are called *Open Access Initiative*. This provide researches with larger amounts of valuable information which can be afterwards analyzed in order to be analyzed for understanding the relations and styles among members of the studied phenomena (Borgatti & Foster, 2003).

3.1.2 Social Network Analysis Theories

SNA is built upon one basic assumption which is any problem can be represented as a social network. This means any problem can be understood throughout their members' relationships, connections and how do they interact with each other. Many well-known theories have been evolved from social network context (Scott, 2000).

There are four basic assumptions on which almost all network theories are based on: individuals or actors are referred to as separate nodes within the formed network, relations are formed from any type of connection or interaction occurred between individuals, individuals can be more authorized through their positions within the network, and building a strong relations among interactive actors in the formulated social network (Wasserman & Faust, 1994).

Here is a listing of the main theories which are based on SNA as their research approach (Albert & Barabási, 2002):

1. Diffusion of Innovations Theory (DIT) (Rogers, 2003),
2. Strength-of-weak-ties (Granovetter, 1973),
3. Small world theory (Travers & Milgram, 1969),
4. Degree distribution based on a power-law distribution (Albert & Barabási, 2002).

Diffusion of innovations theory (DIT) studied the extracted social networks and it explores how social structure of social network affect the speed of information, news, and new ideas distribution. Moreover, investigating the different roles of individuals within the network such as finding out how their behavior based on its position within the network affect the overall output within the network (Giani, 2010).

The *small world theory* is another example of social network based theories. It stated that every node (i.e. individual) within a network is connected to any other nodes within the network through very minimal number of steps. This theory is based on the old well-known concept *six degree of separation*. This concept date back to the experiment done by the American psychologist Milgram (Travers & Milgram, 1969). In his experiment, Milgram asked a sample of people to distribute a specific information, this was done through sending a message through a specific group members. It was found out that the average distance between any two members was six steps (Watts & Strogatz, 1998).

Another theory is the *strength-of-weak-ties* theory. Granoveter stated that the importance of weak ties in receiving information. The formation of sub-groups or cliques is a reflection of having a compatible and similar individuals within the whole network. These groups share similar interests and knows all the information about each other more than the outsider group members. In order to gain a full coverage of the network, the strong relations established within the groups still need support from other neighbor actors represented by the week ties from other groups through connectors.

Furthermore, after the 1990s researchers have developed newer theory based on social network which is *structural holes* developed by Burt. Burt claimed that having more weak connections and relations among the group members reflect in having a high tendency of finding holes i.e. disconnections among the social structure of the represented phenomena. The existence of these holes in the social structure is critical in information distribution among the whole network, especially within the structures that have many holes. These holes represent nodes with what is called gate keeper role that control the communication with different disconnected social structures or subgroup (Lin, 2001).

Scale-free networks is another term describing social network distribution. The distribution is based on power law one. Related to social networks i.e. the social problem under study means a small number of nodes have the larger number of connections while the larger number of nodes have few connections (Barabási & Bonabeau, 2003).

3.1.3 Examples of Social Networks and their Applications

Social networks and SNA are considered a multidisciplinary approach which can be applied in many research disciplines. Consequently, social networks have many different applications in sociology, epidemiology, biology, criminology, and physics. Any phenomena within these research area can be studied from social network perspective as long as this phenomena is represented as a social network (Kadushin, 2005).

Accordingly, SNA is applied to computer science problems like semantic web, social recommendation systems, software development, and collaborative learning. Other examples include representing and analyzing the links among web pages, scientific paper citations, electrical powers grids, etc. (Borgatti *et al.*, 2009)

More specifically, in computer science area, new problems, online relations were studied from a social network perspective. Due to the new ideas and development of Web 2.0 services and sites, endless examples of networks can be found, for example, online shopping sites Examples of network applications are:

- Citation networks: are networks that are formed from the relations between their titles i.e. research field, or the relation between their authors, or their shared references list. Vazquez (Vázquez, 2001) studied network of paper citations.
- Network services network: networks extracted from websites about the relation of friendship, or websites about academic and work connections. For example (Catanese *et al.*, 2012) that studied friendship relations, and (Newman, 2001a,b) who studied collaborative networks.
- Information networks: for example networks of neural networks (Amaral *et al.*, 2000), or phone-call networks (Aiello *et al.*, 2000; Abello *et al.*, 1998).

Furthermore, another application of SNA is representing the Web as a graph. The extracted social networks from the web represent the connections between pages hyperlinks. The web is represented as a network where the nodes are webpages and the links between them represent their hyperlinks i.e. the links are based on their authors. Many SNA measures are useful to study the characteristics the networks, most useful is eigenvector centrality measures like PageRank to find out the most visited pages (Kumar *et al.*, 2002).

Consequently, SNA is considered as an analytical tool to understand the features, characteristics, roles of the studied problem and its participants. As stated earlier, many research fields utilized SNA measures as the basis for their research. Lately, it has been applied to study learning environments and collaborative learning problems. It can be a helpful tool to interpret the relations among students and their roles they took within the class. Also, it helps to understand students' interactive and communicative pattern and how all of that affect their academic performance (Haythornthwaite, 1999).

Chapter 4 provides a study of how SNA is applied to learning problems and learning environments.

3.1.4 SNA Wider Applications Areas

As previously mentioned, SNA measures and applications are utilized to study the social structure of any formed and studied social networks. Social networks represent any phenomena that can be represented as nodes and links among them.

During the last years, SNA applications got an increased interest in wide number of applications and research fields. For instance, SNA applications have been also used in corporation and business management. The focus is to analyze relationships among workers, information sharing among them, and their relations with their managers. This can find new work strategies to enhance the performance and worker efficiency in the workplace. Another interesting field in which SNA has been applied is studying terrorism. Any terrorist incidents are represented as a network and through SNA the networks can be studied to point out the ones with the most fanatic ideas and the overall social structure between the fanatical group members (Ressler, 2006).

Furthermore, SNA was applied in studying epidemiology. Scientists have studied the incidents of any wide spread diseases. Studying these incidents from SNA point of view can give an explanation of how fast they spread among people. A study by (Krebs, 2006) listed a number of very largely expanded SNA applications:

- "*Examine a network of farm animals to analyze how disease spreads from one cow to another,*
- *discover emergent communities of interest amongst faculty at various universities,*
- *reveal cross-border knowledge flows based on research publications,*
- *determine influential journalists and analysts in the IT industry,*
- *unmask the spread of HIV in a prison system,*
- *map an executive's personal network based on email flows,*
- *discover the network of innovators in a regional economy.*"

3.2 Social Network Analysis Approaches

Basically, SNA is studying relationships established between different individuals within specific environments or phenomenons. The environments or phenomenons could be business corporate, computer networks, users of social online network, or any organization. The nodes of the network can be people, webpages, computer, or any other entity while the links between these nodes represent the relationships between them i.e transmitting information, news distribution. SNA approach can be divided into two main categories for analyzing social structure of the social network, they are visual and mathematical analysis of the formed networks (Krebs, 2004).

Formally, it was clarified the two SNA approaches through the visual and mathematical analysis of social relations. Visualizations of social networks are important in such a it provides a representation of how the members are actually connected and subgroups formulated. SNA a representation of show whether the relation exist between all members of a group, or whether subgroup members are interaction

more ore less with the rest of the members of the group. SNA concentrates on studying network structures through two main types: (1) sociograms (or graphs) are graphical representations of the network using nodes and links, (2) relations values or quantities which are the mathematical results of SNA measures. These indicators are quantitative analysis values of social networks such as degree centrality, density, cohesion, etc. (Wasserman & Faust, 1994).

Visualization can be defined as generating images of the social networks along with their nodes and edges. It is considered as an essential supplement to the mathematical analysis of the network. Moreover, the visual representation of social network ease the detection of subgroups and disconnected nodes within them. The graphical representation of the social network is called a *sociogram*. Moreover, these sociograms explores all the various roles the network members play. For example, which node is taking the lead in initiating the communication, the one with very less communication, and the subgroups found (Haythornthwaite, 2002).

On the contrary, the mathematical quantitative analysis of social network is accomplished on three parts of the network. These levels are the whole network, the subgroup i.e. community, and node i.e. actor level. For example, understanding nodes' location within the network i.e. through calculating some measure, can give an understanding of its role i.e. broker, information distributor, or isolated within the group members (Haythornthwaite, 2002).

Network researchers identified a number of fundamentals that are considered the basis for any social network theory. They are summarized as follows:

1. visualizing relationship between individuals i.e. communication between individuals or sending/receiving messages through graphs or sociograms,
2. finding out what are the aspects that affect the established relations between network members,
3. summarizing the results of the social roles of the social networks, like brokers or most influential nodes,
4. applying some suggestions for enhancing the communications and interactions among network members (Dekker, 2002).

There is a definite difference between studying personal information of each node within the network and studying the general social structure of the social network i.e. how the nodes are distributed and formed the social network. Wellman (Wellman, 1988) highlights this concept in such a way: "behavior is interpreted in terms of structural constraints on activity rather than in terms of inner forces within [actors]".

Recently, SNA applications have been applied into another research focus area like studying e-learning problems, for example: (1) Examining new teaching hypothesis and discovering any dependencies among the studied factors, (2) providing an evaluation of students' outcomes and performance, and (3) recommending new learning strategies like introducing group work among students, new enhancing

learning activities, and new methods. All of these are introduced as a result of analyzing the class interactions (Dekker, 2002).

3.2.1 Data Aggregation Methods

Different methods can be employed for collecting data for later analysis from a social network perspective. The earliest methods are: (1) questionnaires, (2) interviews, (3) observations, (4) archival records, or (5) experiments (Krebs, 2004).

During the early SNA applications, the main data collection approaches were questionnaire/survey. Questionnaire based on distributing questions among focused group and gathering their answers and responses. These answers are the main material for analysis. However, this type of data can be erroneous, biased, and require a lot of work. It suits better not very large population size (Knoke & Yang, 2008).

This study does not use the previously mentioned methods for data collection, it is based on data extracted from Moodle database. Further analysis is found in chapter 5.

After collecting the data for analysis, they have to be converted into a format that can be processed by any SNA software applications. The converted data into the proper network format, they are analyzed through SNA application programs like UCINET, Pajek, etc. SNA data is represented using matrices, graphs, and sociograms (Garton *et al.*, 1997).

Basically, there are two main approaches for gathering and collecting data for analysis through social network approach. First, *ego-centric networks*. Ego-centric, also called personal networks, studies specific node's characteristics and properties comparing to the whole network. This approach studies a central nodes and their influence effect against the whole network performance (Zaphiris *et al.*, 2009).

Second, *whole networks* approach which studies the studies the overall general patterns which are generated from the nodes distributions and their connections. This approach aims to study a specific phenomena of individuals relationships, such as email message exchange, Facebook friendship relations, or twitter followers, etc. (Zaphiris *et al.*, 2009).

Moreover, there is another classification of social networks which is related to their types i.e. the nature of the relations and type of individuals being studied. This classification is as follows: *one-mode network* and *two-mode network*. One-mode network is a social network that is composed from one particular type of node. The relation between any two nodes within the network represent one single relationship; the link between node X and Y is a friendship relation or message exchange between them (Sie *et al.*, 2012). One-mode network is the suitable format

for studying social networks extracted from learning environments i.e. learning networks.

On the contrary, the second type of social network is *two-mode network*. This social network is composed of two different types of nodes, like members in a club and events held. Two-mode network is referred to as a membership or affiliation network since the relationships represent a membership in any organization or a marriage relation or family relations (Marin & Wellman, 2009). This type of relation is considered transitive which means it can be converted into one mode network. This ease the analysis process of the social network and applying the normal SNA measures (Sie *et al.*, 2012).

3.2.2 Level of Analysis

The social network can be studied based on different level of analysis. Borgatti and Foster (Borgatti & Foster, 2003) defined three main levels: *node*, *subgroups*, and *network*. The level is chosen based on the focus of the study. Here is a summary the most important measures that are classified based on the level of analysis (Gretzel, 2001):

- Node level measures include centrality measures like degree, betweenness, and closeness. They can point out the different roles of the nodes like isolates or bridges, etc.
- Subgroups measures include measures that detect cliques, cohesive subgroups, and components.
- Network level measures studies the overall general centralization and density of the studied network.

Furthermore, studying different levels of the network provide a wider and full understanding of the social network. The whole network measures studies the number of links forming the network. For example, density measures the number of links within the network compared to the total number of links that can actually be built (Gretzel, 2001).

Subgroup level measures detects the number of subgroups formed from the overall network communications. The connections within the members of these groups are stronger than the members outside in other groups. The members of these subgroups have similar characteristics or behavior styles (Passmore, 2010).

Finally, applying measures on individual level can detect the characteristics that distinguish the main actors with the leading roles. For example, degree centrality measure finds out the most popular individual within the network i.e. one with maximum number of connections. Betweenness centrality also detects the importance of the individual's position in the network to pass the information among the other network members. Finally, closeness measure is related to the minimum distance

of how one node is far away from all the others within the network (Passmore, 2010).

In section 3.6 the mathematical equations of the previous measures are presented.

3.2.3 Representing Social Relations

A relation defines any type of association or tie between any two nodes i.e., they are either object or individuals. This two way relation, also called dyadic relation, can be either direct or indirect relation. In direct relation the node which starts the communication differs from the one receiving it i.e. sending messages. On the contrary, the indirect one the relation direction does not matter i.e. membership relations. Moreover, a tried relation between a group of three nodes. A subgroup is a subset of nodes that share similar relations and disconnected from the rest of the network. Finally, a social network is a number of connected individuals that are grouped through one distinct relationship (Passmore, 2010; Gretzel, 2001).

The output format that represents the results of social network measures are either sociograms or matrices. A sociogram is a graphical visualization of the individuals and their connections through nodes and edges. The indirect relation is represented with an edge without an arrow. While the directed relation is represented with an edge with arrows i.e. distinguishing the sender from the receiver (Passmore, 2010).

The second format is an adjacency matrix or sociomatrix which is a algebraic translation of the network through its nodes and links. Basically, a sociomatrix is an array of entries arranged in rows and columns. The entries within the array are defined through dimensions of N^2. A specific array entry or a cell is defined by X_{ij}. For example, the entry X_{12} represent the entry input in the first row and the second column respectively (Knoke & Yang, 2008; Wasserman & Faust, 1994)

The mathematical representation through sociomatrix can be both symmetric or asymmetric. For example, an entry in a symmetric matrix means the entry in X_{23} is similar to the one in X_{32}. This represent an indirect relation. Whereas in an asymmetric matrix the entry in $X_{23} = 0$ does not equal to $X_{32} = 0$. This represent the directed relation where the rows represent and columns represent different set of actors regarding to their direction (Knoke & Yang, 2008; Wasserman & Faust, 1994).

The previously defined terms are more interrelated with graph theory. More in depth analysis is provided in the following sections and subsections.

This figure is translated into a matrix as follows: $\mathrm{M} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

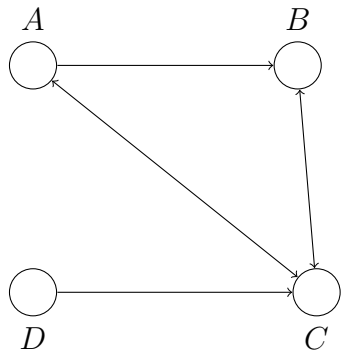


Figure 3.1: Simple example

3.3 Graph Theory Concepts and Basics

This section is a brief introduction to graph theory, network theory, and basic concepts in graph theory. Finally, it explains the SNA measures used within this study.

3.3.1 Introduction

SNA as a research approach has its origin in graph theory especially benefiting from mathematical and structural approaches. Social networks represent the connections and relationships among nodes of different types either people, organizations, webpages, scientific papers, or computers. Moreover, these networks are used to measure, analyze and the different types of connections between its members (Krebs, 2006).

Basically, all SNA measures have their origin in graph theory and allow analyzing social networks through mathematical and visual analysis of the relations among networks. These measures allow to study the network characteristics and the structure of the networks. It also studies the roles of network members and their behavior (Passmore, 2010).

As stated earlier, many of the terminologies and measures of SNA have their origin in graph theory. This approach help to analyze the data quantitatively through the measures after representing the data as network. Utilizing graph theory approach, a network is defined as a graph G. A network represented as a graph G such that $G = (V, E)$ which is a basic illustration of the nodes V and their links E (Wasserman & Faust, 1994).

The number of nodes are composing the network are V and their edges which define the cardinality of the graph and indicated by $|V|$ and $|E|$. Any two adjacent nodes are connected by a link e which represent the type of the relation studied in the

network. The adjacent nodes are referred to as neighbors (Wasserman & Faust, 1994).

3.3.2 Graph Theory Definition

Graph theory along with its applications, theories, and mathematical representations, have been applied in various research fields like psychology, communication, business applications, and many others. Moreover, it has a large application and basis in SNA. Graph theory support SNA through representing the nodes graphically and their structure of relations. Graph theory also introduces the basis for mathematical measures and equations that allow quantitative analysis of the extracted social network (Knoke & Yang, 2008).

A graph representation is the simplest way to represent the structure of the social relations. It is considered as a simple graph or image of how actually individuals or nodes are connected. It allows also discovering new connections or relations that are hidden (Martino & Spoto, 2006).

Furthermore, graph theory has a large applications in computer science. SNA, originated from graph theory, has been largely used in computer science problems. SNA from a computer science perspective is considered as an applied approach of graph theory. It studies the formulated network from both structural and mathematical aspects (Wasserman & Galaskiewicz, 1994).

The mathematical representation of a graph G can be described as an ordered pair such as $G = (V, E)$. The nodes that are composing the network are $V = \{v_1, \cdots, v_n\}$ where the links are $E = \{e_1, \cdots, e_m\} \subset \{\{x, y\} \mid x, y \in V\}$ (Passmore, 2010).

All the previously mathematical definitions of the graphs can be also represented as graphically as in figure 3.2. This figure shows the simple basic structure of any graph or social network along with their nodes and edges:

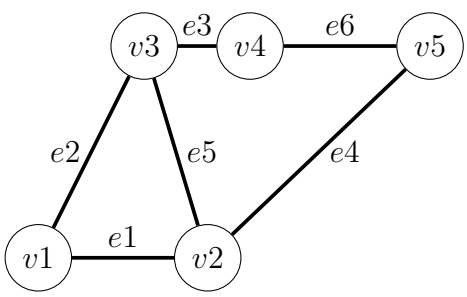


Figure 3.2: Basic graph with nodes and their links, adapted from (Passmore, 2010)

3.3.2.1 Graph Types

The previous subsection 3.3.2 explained the concept of graph theory, its origin in SNA, and the mathematical definition of graphs. This section shows basic graph types defined in the literature.

As mentioned earlier, a graph can be represented as $G(V, E)$ with its nodes and edges. The graph size is defined as the number of nodes composing the graph, such as $|V|$ (Wasserman & Faust, 1994).

What follows is the list of the main types of graphs with their mathematical definitions (Wasserman & Faust, 1994):

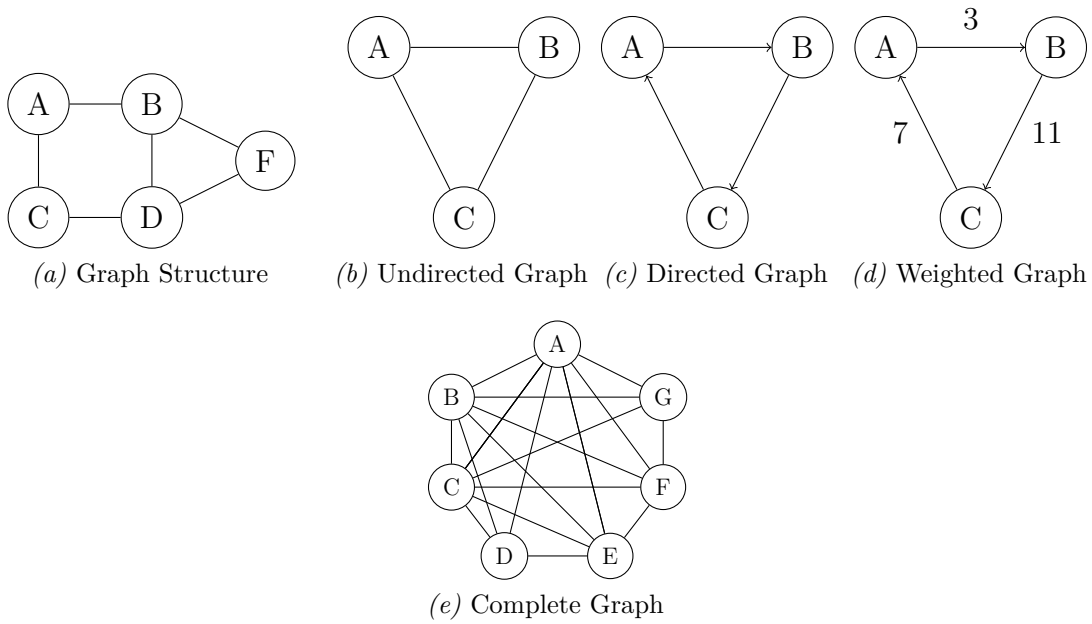
1. Basic graph which describe a simple fundamental structure, figure 3.3a.
2. Indirect graph describes the symmetric relations where there is no difference between the direction of the communication, figure 3.3b.
3. Directed graph i.e.e diagraph describes the asymmetric relations where it is important to distinguish the sender and receiver nodes, figure 3.3c.
4. Weighted graph assigns a cost or a strength to the relation, figure 3.3d.
5. Sub-graph describes a subset of the whole graph that are tightly connected.
6. Complete graph is shown in figure 3.3e.

An undirected graph is defined as follows: a node u is adjacent to (or is direct neighbor to) v if $\{u, v\}$ is an edge included in the set E. This edge is denoted as $e(u, v)$. These nodes u and v are referred to as endpoints of $e(u, v)$, and it is said that the edge $e(u, v)$ is incident with vertices u and v. All edges in an undirected graph are symmetric (Wasserman & Faust, 1994)

3.4 Network Theory

Network theory is a research approach that is originated in graph theory and have a large number of applications in computer science. Its applications varies in many areas like physics, biology, economics operation research, sociology, and many others. Basically, its focus on analyzing the graphical representation of social networks i.e. the social relation represent different types of relations. Applications of network theory in computer science include analyzing websites and their users' interactions and online social networks. Moreover, SNA as an application of network theory analyzes and studies social networks from both mathematical and visual perspective. First, the mathematical analysis done through many indices to study the network from a quantitative perspective. Second, visualization and graphical representation of the network by sociograms which represents how nodes are connected and their relations are established (Martino & Spoto, 2006).

The distinction between a graph and a network can be defined mathematically as follows. Previously in subsection 3.3.2, a graph is defined as: $G(V, E)$, where V is



(a) Graph Structure *(b)* Undirected Graph *(c)* Directed Graph *(d)* Weighted Graph

(e) Complete Graph

Figure 3.3: Different Graph Types, adopted from (Passmore, 2010)

the set of vertices and E is the set of edges connecting them. On the contrary, a network represented as N is defined as $N = (U, L, F_U, F_L)$, this network is composed of a graph $G = (U, L)$, which is an ordered pair of a vertex set U and a line set L, extended with a function F_U specifying a vector of properties of the units $(f : U \to X)$ and a function F_L specifying a vector properties of the lines $(f : L \to Y)$. The set of lines L may be viewed as the union of a set of undirected edges E and a set of directed arcs A $(L = E \cup A)$. Each element e of E is an unordered pair of units u and v (vertices) from U, that is, $e\,(u : v)$, and each element a of A (each arc) is an ordered pair of units u and v (vertices) from U, that is, $a\,(u : v)$ (de Nooy, 2009).

Basically, a network is a graph but it has more characteristics of their composed nodes. These characteristics represent the additional information about the nodes, regarding their type like age, name, and personal interests. These characteristics are classified as either discrete or continuous characteristics. These characteristics can be represented by assigning weights to the edges i.e. relation between two nodes (de Nooy, 2009).

Additionally, SNA is seen as the practical approach of graph theory. SNA take high advantage of analyzing the underlying patterns of the graphs and their mathematical characteristics of their nodes (Wasserman & Galaskiewicz, 1994).

3.4.1 Network's Types

A network structure is the graphical representation of the links which mark the existence of any relationship connecting a set of members i.e. students, workers, colleagues within any specific context. Whereas the social structure is basically a network structure but it include the characteristics of the network members such as pointing out the different groups found in the whole structure along with various types of nodes i.e. with distinct characteristics (Coulon, 2005).

Fundamentally, any defined social structure is composed of these defined models: (1) blockmodels, (2) hierarchical clustering, (3) multidimensional scaling. These models are proposed earlier in literature but have gained larger attention with the large usage of SNA. Their main function is to get information about social relation and structure of the network. The social structure are extracted from a relational database through grouping similar shared nodes and the result of SNA measures. Also, social networks are considered a visual translation of the various social roles of nodes composing it. The overall patterns represent the hidden patterns and relations in the social network (Watts, 2004).

The following subsections define and explain the meaning of one-mode and two-mode networks.

One-mode Network

Basically, the type of the social network is determined by its mode. A mode of a network represent the number of variables that are building the social network. One-mode and two-mode networks are the largest used types. However, one-mode networks are more used than two mode and the two-mode networks must be converted into one-mode to be further analyzed Passmore (2010). The following subsection 3.4.1 discusses the two-mode network basis.

One-mode networks are composed of only one single type of nodes. These nodes can represent either people, events, organizations, webpages, or computers. Most importantly, the link between the any two nodes represent a single type of relation. Relations in this types of networks are studied on node level; the characteristics of the individual nodes composing them (Passmore, 2010).

For example, studying friendship relations is done through one-mode network, like studying friendship relations of students in school. However, studying two-mode network is studying the relation between two categories of individuals' relation, like workers and their organizations (Wasserman & Faust, 1994).

Two-mode Network

Two-mode network represent a type of networks that are formulated from two different node types. The relation in two-mode network is created from connection between different types of actors. For example, it can represent a membership or marriage relation (Wasserman & Faust, 1994).

Moreover, relations within two-mode networks are also called affiliation. Affiliation networks are composed of two types of nodes so that one type is affiliated or connected to the second type. There is two main types of node sets one is actors and the second is the defined events. Some examples of two-mode networks are: (1) colleagues and membership in organizations. (2) authors, papers, cited papers. (3) people, occasions, attending. (4) people, marriage relationships, etc. (de Nooy *et al.*, 2005).

The mathematical definition of two-mode network is described as (U, V, R). In this definition U and V belong to two discrete groups of nodes. The relation is R is defined as $R \subseteq U \times V$. Mainly, there is no weight assigned to the relation in this type of networks. The majority of SNA measures are applicable to on one-mode network, therefore two-mode networks have to be converted into one-mode network without changing their relations (de Nooy, 2009).

Figure 3.4 shows a simple two-mode network where the squares represent papers and the circles represent their authors.

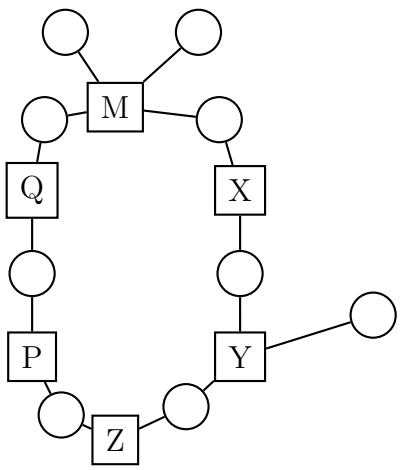


Figure 3.4: Simple Example of two-mode Network, adapted from (Passmore, 2010)

3.5 Elementary Concepts in Graph Theory

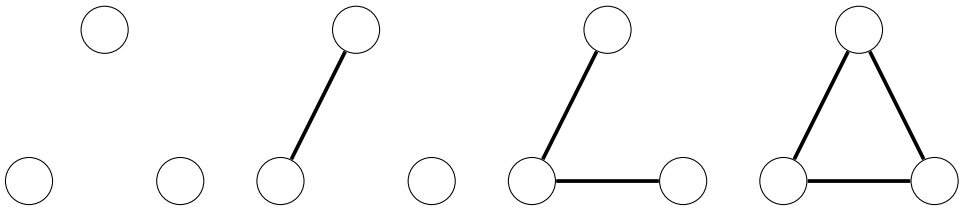
This section studies graph theory concepts and their origin in SNA measures along with their mathematical definitions. It also shows the relation between graph theory and SNA measures.

Throughout section 3.3.2, it provided the mathematical definition of a graph G as $G = (V, E)$. It also showed graph types in the literature. The following subsections provide definitions and meaning of the basic concepts of graph theory.

Dyad and Triad

A dyad represent a relation between two nodes or actors. Different network measures are associated with dyad relation. In undirected graph the definition of dyad relation is limited of being connected or not only. Dyad relation is considered the basic relation. In a graph with unordered pair of nodes such as a social network, only two states of dyads are considered the basic item of analysis in SNA.

On the other hand, a subgroup of three nodes and all the possible connected relations is called a triad. Basically, among a triad subgraph, there can be four different relations formulated, they are shown in figure 3.5.



(a) 0 Lines *(b)* 1 Lines *(c)* 2 Lines *(d)* 3 Lines

Figure 3.5: Example of potential connection from a triad (Wasserman & Faust, 1994)

Path

Another important term is the path between any two nodes. It is defined as the sequence of nodes that connection two main nodes. All the nodes and connections that are forming the path between the starting and end node should be different. The path length is calculated as the number of of nodes composing it. If the path contains similar nodes, then it is called trail. If the nodes or the connections are not different; no repetition then the path is called a walk. More specifically, path measures if two nodes are connected and find out its distance. Path measure is considered an essential component in other measures like geodesic distance, diameter and centrality measure (Wasserman & Faust, 1994). These measures are illustrated in more depth in the following sections 3.6.3.

There are other measures that are based on path in its calculations, such as cycle and shortest path. A cycle is a special case of a path where the starting and final nodes are identical. Whereas the shortest path is the one with the shortest number of nodes composing it. It is referred to as geodesic distance (Wasserman & Faust, 1994).

Geodesic Distance

Geodesic distance is defined as the shortest path that is connecting any two nodes in the graph. This definition differs regarding directed or undirected graph. For directed graph, the geodesic distance for two nodes is the shortest path between any pair of nodes in a graph. In a directed graph, the pair of nodes have different geodesic distance regarding the direction of their connections (Wasserman & Faust, 1994).

The following are two equations that represent calculating the geodesic distance between any two nodes (Borgatti, 1995):

$$L = \frac{1}{\frac{1}{2}n\left(n+1\right)} \sum_{i \geq j} d_{ij} \tag{3.1}$$

Second definition: 3.2 (Borgatti, 1995):

$$L^{-1} = \frac{1}{\frac{1}{2}n\left(n+1\right)} \sum_{i \geq j} d_{ij}^{-1} \tag{3.2}$$

The first definition of L is used for graphs or networks that are composed of disconnected components. The second equation represent the average distance i.e. reciprocal distance of all pairs composing the network (Wasserman & Faust, 1994).

Subgroup and Group

A subgroup is any subset of nodes that are highly connected together seen as a component of the whole network. Studying and detecting subgroups are essential in SNA measures and network analysis. The difference between subgroup and group is related to the number of nodes formulating each one. The members of either group or subgroup have one shared characteristics (Wasserman & Faust, 1994).

The mathematical definition of a subgroup extracted from a graph $G\left(V, E\right)$ is defined as $R\left(X, Y\right)$ where its nodes are subset of the original graph G. In any subgroup there is one specific node that is called boundary node which is considered the connection point between this subgroup and all the other nodes in the whole graph. Moreover, it can be calculated a boundary size for each subgroup which is the summation of the boundary nodes. Another important term description the nature of connection within the graph is the cut-edge. A link is called a cut-edge if its two edges are members of two different subgroups (Wasserman & Faust, 1994).

Density

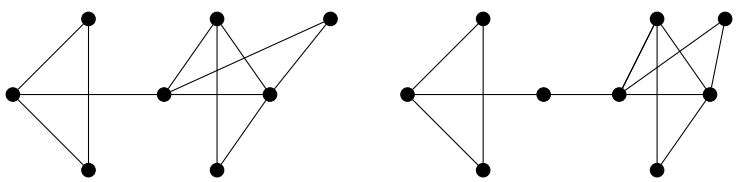
Density of a graph is defined by the relation of summation of all links that are actually happening within the graph to all possible links that might occur within the graph. The density values of the graph is highly related to the size of the graph. Larger densities values belong to larger graph sizes. In order to compare the densities of different graphs, they have to be of the same size to be compared (Wasserman & Faust, 1994).

Graph's Connectivity

Basically, graph's connectivity illustrate a measure of graph if it remains connected even nodes or connections are removed from the whole graph. Graph connectivity is related to centrality measures that are used to analyze any social network on the node level. Centrality measure points out the nodes with the most important roles within the network (Wasserman & Faust, 1994). Centrality measures are explained in section 3.6.3.

One of the important concept related to graph connectivity is cutpoint. Figure 3.6a represents a graph that is divided into a number of separate components once some nodes are taken out from it. A node is defined as a cutpoint after finding out the number of components of the graph containing this node is smaller than those after removing it from the graph. Correspondingly, a number of nodes are considered a cutpoints because their removal from the whole graph make the graph to be divided into a number of disconnected components (Wasserman & Faust, 1994).

Accordingly, some of the found cutpoints are essential in the overall connectivity of the graph in this case they are called bridges. Similar to cutpoints, a number bridges that are found make the graph divided into a disconnected components. This is represented in figure 3.6b. Graph connectivity, cutpoints, and bridges are all related to graph cohesiveness (Wasserman & Faust, 1994).



(a) Example of cutpoint *(b)* Example of a bridge

Figure 3.6: Example of cutpoints and bridges (Wasserman & Faust, 1994)

3.5.1 Components

A component is a subgroup or a subset of the graph that are tightly connected with each other but separated from the reset of the graph. The separated i.e. disconnected components have problems of receiving news and updates from the rest of the network or graph. Basically, for directed relations in a network, there are two types of components weak and strong components. A weak component is a group of nodes that are connected together apart from the whole graph but their relation does not take into account the direction of the connection, most important its existence. Whereas forming a strong component require existence of directed path between two nodes so that they belong to the same component (Hanneman & Riddle, 2005).

3.6 SNA Measures

This section discusses the methods for analyzing social networks, specifically the ones that are used within this study and used to analyze learning networks within learning environment. Throughout the following subsections, mathematical equations and figures are presented to illustrate each one of the concepts. This section discusses the following measures: *density*, *centrality*, *cohesiveness*, and *clustering*.

3.6.1 Network Size

Network size is a measure that is calculated depending on the type of the relation of the network. In principle, a network size N is the number of nodes composing the network. It is important to understand the overall network structure and the composed social relations (Hanneman & Riddle, 2005).

Within networks of smaller sizes, the tendency of finding disconnected nodes is minimal since nodes are more likely to be connected to their neighbors. On the contrary, networks of larger sizes may have a number of disconnected nodes since it may be difficult that all the nodes are connected to all the others. Calculating the network size differs from directed and undirected network. For a directed network the network size is the number of all possible connections $n * (n - 1)$. Whereas in the undirected network it is calculated by $n * (n - 1) / 2$ (Hanneman & Riddle, 2005).

Moreover, the degree of each actor in the network represent the maximum number of connections it can have. This is a valuable measure about the network since it provide information about the patterns of how nodes are connected (Hanneman & Riddle, 2005).

3.6.2 Density

Density measures the how highly connected a graph or network is. In special cases where all the nodes are connected to each other, this case is called a complete graph or complete network. Mainly, density is a relation between the number of links actually occurred to number of possible links that could really occur (Hanneman & Riddle, 2005).

Network density, referred to as δ_{density}, is a measure that expresses the general connectedness among network members. It is a percentage of the number of connections that are actually happening in the network to the overall number of links that could possibly occur within the network (Hanneman & Riddle, 2005). Equation 3.3 measures the network density δ_{density} (Wasserman & Faust, 1994)..

$$\delta_{\text{density}} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} x_{ijk}}{n \cdot (n-1)} \quad \text{where} \quad i \neq j \neq s \tag{3.3}$$

where $n \cdot (n-1)$ is the total number of ties possible and n is the relation under study (Wasserman & Faust, 1994).

Moreover, the density measure is a representation of the coherence of network's members and their relations. It gives an indication of how fast the news are distributed among all the network members. This can be referred to as the speed of information distribution (Knoke & Yang, 2008).

A network with N number of nodes, the number of possible of pairs regardless of the direction can be calculated as $\binom{N}{2} = N(N-1)/2$. On the contrary, for a network with directed relations, the largest possible number of ordered pairs are $N(N-1)$ (Wasserman & Faust, 1994).

Moreover, for undirected graph or network the density, δ_{density}, is calculated as the percentage of the number of links within the network (L) to the largest number of links that can actually happen within this network. This is presented within this equation 3.4 (Wasserman & Faust, 1994):

$$\delta_{\text{density}} = \frac{L}{g(g-1)/2} = \frac{2L}{g(g-1)} \tag{3.4}$$

The density measure for directed graph or directed relation network is computed based on this equation 3.5 (Wasserman & Faust, 1994):

$$\delta_{\text{density}} = \frac{L}{g(g-1)} \tag{3.5}$$

Mainly, the results of density measure ranges from 0 where there is no connections in the network, to 0 where this is a complete network or graph (Hanneman & Riddle, 2005).

3.6.3 Centrality Measures

Centrality measures analyzes the network based on node level and investigates the prominent roles played by the network members. This measure gives an understanding of the social roles of the most important actors within the network as well as how these nodes form the social network pattern (Passmore, 2010).

There are four different centrality measures defined in order to study different roles of network members. They are: degree centrality, closeness centrality, betweenness centrality, and eigenvector centrality.

3.6.3.1 Degree Centrality

Degree centrality can be measured for each single node of the network. It represents the number of links each node have. Mathematically, degree centrality $d\,(i)$ for a single node i is defined as (Wasserman & Faust, 1994):

$$d(i) = \sum_j m_{ij} \tag{3.6}$$

where $m_{ij} = 1$ if there is a link between nodes i and j, and $m_{ij} = 0$ otherwise. This measure can be standardized by dividing over $(N-1)$. The final equation is $d_s(i) = d(i)/N(1)$ (Wasserman & Faust, 1994).

On the other hand, degree measure $C_D\,(n_i)$ for nodes in an undirected network, is simply computed as the summation of all the links each node has.

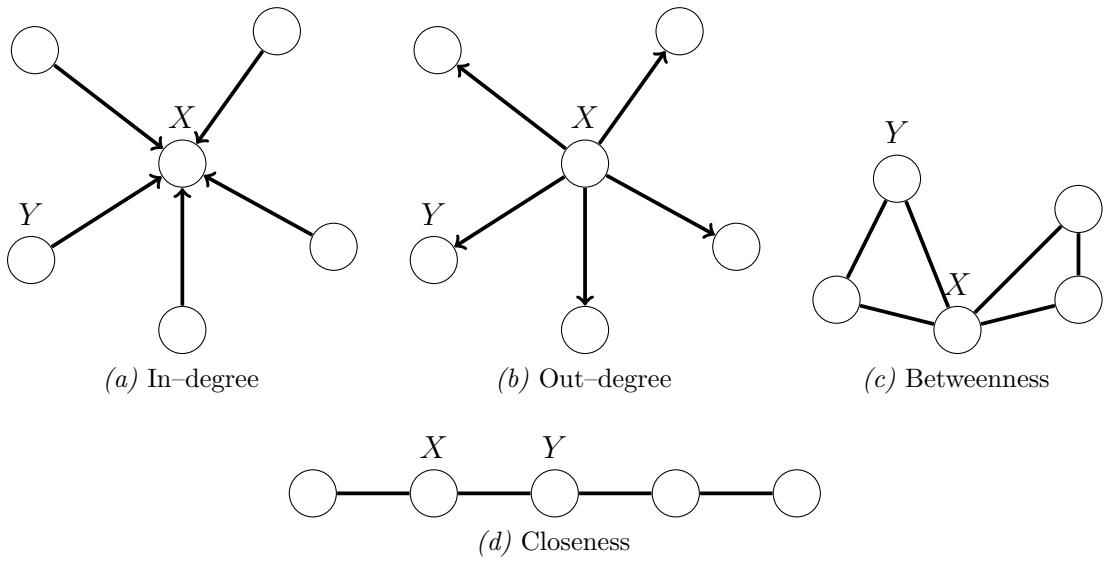
The degree centrality of an undirected network, $C_D\,(n_i)$, is calculated by the sum of all the ties of actor n_i to the other actors of the network. In order to compare two networks of different sizes based on the centrality measure, this measure must be centralized as well through standardized version by dividing the centrality C_D over the largest number of connections $(n-1)$ (Wasserman & Faust, 1994):

$$C_D\,(n_i) = \frac{\sum_j x_{ij}}{(n-1)} = \frac{\sum_j x_{ji}}{(n-1)} = \frac{C_D}{(n-1)} \quad \text{where} \quad i \neq j \tag{3.7}$$

For directed graphs, degree centrality measure is calculated differently. Since it must take into account the differences in the direction of the relation. Moreover, the degree centrality measure helps in finding out the the most influential node in the network. For directed relation, degree measure distinguish the difference between node with

high power i.e. having the largest number of outgoing links, or the nodes with the highest influence i.e. having the largest number of incoming links. Mainly, the nodes with high degree measure values have many favored advantages due to its position within the network (Hanneman & Riddle, 2005).

Finally, degree measure is an important measure basically in larger number of networks. However, its drawback resides that it takes into consideration of direct ties and this eliminates the importance of node position within the network (Hanneman & Riddle, 2005).



(a) In–degree *(b)* Out–degree *(c)* Betweenness

(d) Closeness

Figure 3.7: Centrality measures illustration. Adopted from (Wasserman & Faust, 1994)

3.6.3.2 Closeness Centrality

The second type of centrality is closeness centrality. It is defined for a single node as the summation of the overall distant from this node to all the other nodes within the network. The mathematical definition of closeness centrality $c(i)$ of node i (Wasserman & Faust, 1994):

$$c(i) = \sum_{i} d_{ij} \tag{3.8}$$

In the equation, d_{ij} presents the summation of links that exists in the shortest path between two nodes i to node j. The larger output values of closeness measure, the less central this node within the network. Moreover, the standardized closeness values is formulated as $c_s(i) = (N-1)/c(i)$ in order to remove that extreme values (Wasserman & Faust, 1994).

When calculating the closeness measure, it takes into account both direct and indirect links of the nodes within the network. This measure studies not only the direct links of a node to all of its neighbors but also to all the other nodes within the network. Basically, closeness centrality finds out how much the node is placed in the shortest path that is calculated to all the other nodes within the network. This measure is only calculated for connected networks (Wasserman & Faust, 1994).

Furthermore, this type of measure is related to the speed of information distribution between nodes of the network. The node with smaller closer values indicates its preference in reaching information much more faster than other nodes. This make the removal of this node from the network is critical in sending and the speed of information distribution to all the other nodes in the network (Hanneman & Riddle, 2005).

3.6.3.3 Betweenness Centrality

Betweenness centrality is defined by the indication of how many times this specific node lies on the shortest path which pass through the node. It is also an indication of the number of occurrences where this specific node is used by the network members to distribute the information (Hanneman & Riddle, 2005). The mathematical definition of betweenness centrality $b(i)$ of node i is as follows (Wasserman & Faust, 1994):

$$b(i) = \sum_{j,k} \frac{g_{jik}}{g_{jk}} \tag{3.9}$$

The equation g_{jk} represents the summation of the shortest path from node j to the other node $k(j, k \neq i)$. Also, this value g_{jik} represents the summation of shortest path from node j to node k which is measured through node i. The standardized version of this measure is calculated as (Borgatti, 1995):

$$b_s(i) = \frac{2b(i)}{N^2 - 3N + 2} \tag{3.10}$$

The overall betweenness measure C_{network} for the whole network is calculate as the following equation (Borgatti, 1995):

$$C_{\text{network}} = \frac{\sum_j (C_{max} - C_j)}{X} \qquad X \text{ is max value possible} \tag{3.11}$$

In other words, betweenness centrality C_B is the relation of the links that happening between two nodes k and j and that are created through node i. Furthermore, the betweenness of any pair of nodes b_{jk} is calculated by the summation of shortest paths between node j and node k that goes through node i and this is divided

by overall summation of the shortest paths g_{jk} occurring between j and k. The previous calculations are computed for each single node within the network. This measure computes how many times this specific node appear in the shortest path between other two nodes. Finally, having a high betweenness values does not mean to have larger number of direct links, it is only related to the occurrence of this node in the shortest path between other network members (Wasserman & Faust, 1994).

3.6.3.4 Eigenvector Centrality

The final type of centrality is eigenvector centrality. This centrality computes the importance or dominance of a node among the network. This measure also allocates an importance value for each node for nodes with high number of connections and fewer values for those with less connections. A well-known example is Google's PageRank (Passmore, 2010).

Based on eigenvector centrality measure, a node is considered important if it is surrounded by many important nodes in its neighborhood. If a node has many important nodes, then it is important as well (Passmore, 2010). The eigenvector centrality v_i of a node i is calculated as 3.12 (Wasserman & Faust, 1994):

$$C_E\left(v_i\right) \alpha \sum_{v_j \in N_i} A_{ij} C_E\left(v_i\right) \tag{3.12}$$

From the equation N_i represent the neighborhood surrounding the studied node $v_i, \cdots, A_x$ (Wasserman & Faust, 1994).

3.6.4 Connectivity Measures of Network

Connectivity measures of a network is classified as the following (Wasserman & Faust, 1994):

1. clustering coefficient which is the tendency find clusters or subgroups within the network 3.6.5,
2. cohesive subgroups within the network 3.6.7.

3.6.5 Clustering Coefficient

The first measure regarding network connectivity is clustering coefficient. This measure computes the connectedness of a network or the degree of which two nodes close to each to be connected as well. This measure can be calculated for each single node and for the whole network. The clustering coefficient C_i for a node v_i is computed as the summation of the links this node it has with all of its neighbors

and this is divided by the the largest possible number of links that could actually occur (Passmore, 2010). The following equation 3.13 represent how this measure could be computed (Borgatti, 1995):

$$C_i = \frac{|\{e_{jk}\}|}{k_i\,(k_i - 1)} : v_j, v_k \epsilon N_i,\ e_{jk} \epsilon E \tag{3.13}$$

In this equation, a node v_i has a neighborhood N which is defined as $N_i = \{v_j : e_{ij} \epsilon E \Lambda e_{ji} \epsilon E\}$. Also, $k_i\,(k_i - 1)$ represent the number of links this node have with all the nodes within its neighborhood (Wasserman & Faust, 1994).

Social networks that represent the social relations and communication between different individuals mainly have high clustering coefficient values. This considered as indication of high network density. Moreover, high values of this measure indicate a highly connected and thus important neighborhood. In other words, this is related to the high speed of information transmission among the network (Hanneman & Riddle, 2005).

Furthermore, clustering coefficient measure is an indication of cliquishness of nodes of networks. It is considered a measure for connectivity among node members within a single neighborhood. It is a relation of how much two nodes will be connected within a neighborhood chosen without any other prior criteria Hanneman & Riddle (2005).

3.6.6 Network Clustering

A social network can be partitioned into smaller subgroups also called cliques or neighborhoods. The nodes belong to these subgroups share one common characteristics from the whole social relation. The members of these subgroups are highly connected with each other comparing to the other nodes outside of the subgroup or clique. One node can belong to two different cliques at the same time Hanneman & Riddle (2005).

Based on graph theory, a clique is defined as subgroup of network which contains the largest number of links that are happening between the nodes of the clique. There is high correlation between finding cliques and clustering coefficient of nodes of subgroups. The clustering coefficient of single node is computed as the relational between links that are between nodes of the defined neighborhood to the largest number of possible node that could occur between them. On the contrary, the clustering coefficient for undirected network is the same but multiplied by 2 (Hanneman & Riddle, 2005).

Equation 3.14 shows the clustering coefficient of a single node i. Whereas equation 3.15 shows the clustering coefficient for the whole network which is computed

as the average of the clustering coefficients of every single node within the network (Wasserman & Faust, 1994):

$$C_i = \frac{2 \sum x_{jk}}{N_i \left(N_i - 1\right)} \tag{3.14}$$

$$C_{\text{Network}} = \frac{1}{2} \sum_i C_i \tag{3.15}$$

3.6.7 Cohesive

Basically, individuals tend to work together within groups. This is referred to the indication of forming cohesive subgroups. In SNA terminology, these cohesive subgroups are called cliques. The nodes that are forming any defined clique are similar and share similar characteristics, and most importantly they exchange news among them faster than the rest of the network. The denser the cliques, the more faster the information is distributed among them and largely similar nodes either in their beliefs or ideas (Knoke & Yang, 2008).

There are many variations for studying and defining cliques within social networks. One important approach is N-clique which is defined as a group of nodes where there is a link between all the nodes in the clique. Mainly, a clique is defined for undirected relations or once forming a clique the direction of relations are not taken into account. In order to form a clique, only the existence of a relation between any two pair is considered and all the other link characteristics are neglected. A clique S can be found in another larger clique $S^{'}$ under the following condition (Passmore, 2010):

$$S \subset S^{'}$$

The size of a clique is defined by the number of nodes composing it such as $|S|$. The largest clique is the one with the largest number of nodes and denoted as S^{*} (Wasserman & Faust, 1994).

Studying and detecting cliques within social networks provide valuable information about the whole network and their members. It is also form the basis for larger analytical approaches i.e. finding communities within social networks (Passmore, 2010).

Additionally, the cohesion measure of the network express the overall connectedness of the network. The network with high cohesion represents a high tendency of any two actors to be connected within the network (Wasserman & Faust, 1994). Equation 3.16 shows how cliques can be calculated (Wasserman & Faust, 1994).

$$\delta_{\text{cohesion}} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} (x_{ij} + x_{ji})}{\frac{n \cdot (n-1)}{2}} \tag{3.16}$$

Small world phenomenon is considered a theory that is highly related to connectivity, cliques, and clustering. A network formed according to this phenomenon is represented as one group that is highly connected but there is no central nodes (Hanneman & Riddle, 2005).

Figure 3.8 represent a graph and help to understand the clique concept.This figure is presented in (Knoke & Yang, 2008). This figure shows two cliques ABE and $BCDE$.

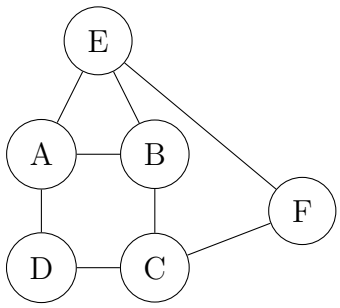


Figure 3.8: Clique Example

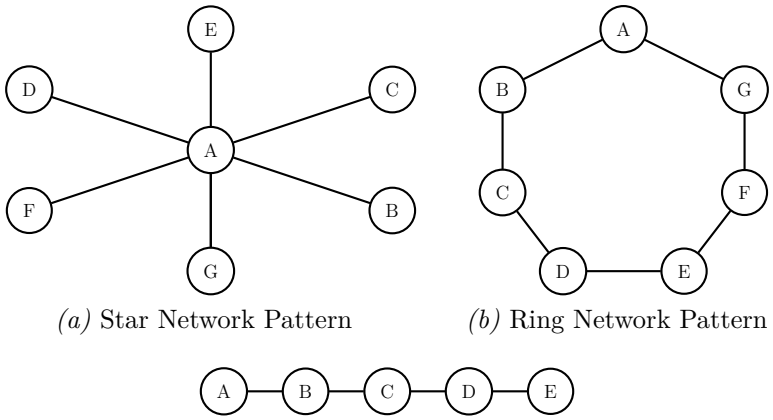
3.6.8 Network Patterns

SNA measures allow studying and analyzing social networks both mathematically and visually. The mathematical analysis is done through the SNA measures that are previously define. On the contrary, visual analysis of social network is done through graphs or sociograms. Sociograms represents how nodes are distributed within the network and connected with each other. Mainly, there are some basic network patterns that are the basic patterns for social networks. These basic patterns are presented in figure 3.9: line network 3.9c, star network 3.9a, and circle network 3.9b (Hanneman & Riddle, 2005).

Here are the definitions for each of the defined patterns (Hanneman & Riddle, 2005):

- A *star* network pattern network. There is one node in the center of the network which having the largest centrality values and all the rest of the nodes are connected only to this centralized node. This network has the maximum centralization measure value.
- A *circle* network pattern. All the nodes are connected in a circular way. All the nodes have an equal centrality measure values. The network centralization of the network is 0.

- A *line* network pattern. All the nodes are connected in a linear sequential fashion. The nodes on the edges having the highest degree centrality values, whereas the nodes in the middle are having the highest closeness and betweenness values.



(a) Star Network Pattern *(b)* Ring Network Pattern

(c) Line Network Pattern

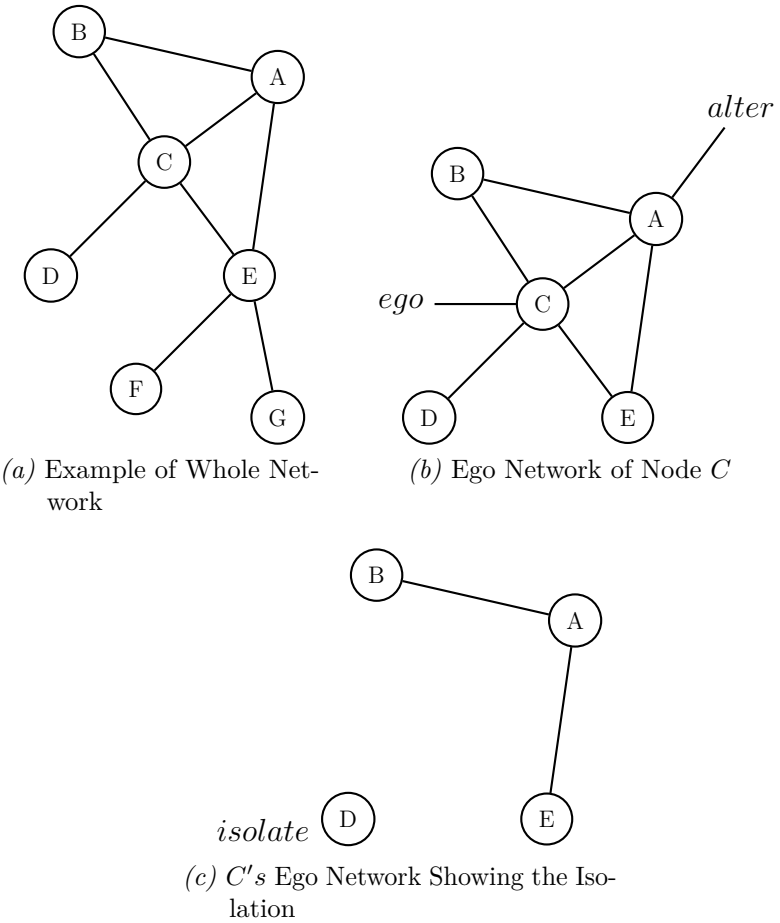
Figure 3.9: Network Patterns

Moreover, there are other terminology related to network patters and how the network are connected; ego network and whole complete network. Ego network is a special pattern that contains the ego network i.e. the studied node along with all its connections and the nodes connected to them. The difference between these pattern types is related to the level of analysis and the objective of the analysis (Hanneman & Riddle, 2005).

Another classification of network patterns is ego-network representation. Ego-network represent a single node i.e. ego and all of its connections and its immediate neighborhood. The nodes within the ego's immediate neighborhood are called alters. Ego-network can be seen as a subgraph of the whole network (Marsden, 2002). Figure 3.10b shows an illustration of ego-network.

However, the whole network represents all the nodes and all their connections. Figure 3.10a is an example of whole network. This figure illustrate the differences between ego and alters and how they can be extracted from the whole network.

Basically, any node can be selected as an ego to be further analyzed. Based on this figure 3.10b. Node C is extracted from the whole network to be studied as an ego-network. This network is formed from C and its immediate connections, they are A, B, D, E the alters of node C. Moreover, in the ego-network the distance between the ego and all of its immediate alters is 1. The ego-network can be defined more broadly through extracting the ego's alter within a distance of k (Wasserman & Faust, 1994).



(a) Example of Whole Network

(b) Ego Network of Node C

(c) $C's$ Ego Network Showing the Isolation

Figure 3.10: Example of whole and ego network

3.6.9 Additional Definitions

This section provides additional definitions of another important SNA measures. They are related to understanding overall social network structure, how the nodes are connected, and how fast information can be reached from one node to all the network members.

Network Centralization

Network centralization is an overall degree centrality measure for the degree centrality for the whole network. Centrality measures are considered important since they define the social power of each node and the social power of the whole network. For a network with directed relation there are two types of network centralization; in-degree and out-degree centralization. The output of degree centralization values ranges from 0% to 100% (Daradoumis *et al.*, 2004).

On the contrary, network centralization is calculated by finding the relation between the difference between the number of links each node has to the largest summation of differences. A network with largest centralization values are the ones where almost all of its links are centralized around one or few single nodes. In other words, the network is centralized around one central node which dominate all network communications (Hanneman & Riddle, 2005).

Centralization measure finds out if the network is centralized around one single node. Density measure on the other hand computes the extent of connectivity among the network. Both centralization and density measures are considered supportive and important ones to understand the overall social structure of the network. The denser the network, the higher the higher the chances are to find one central node in the network (Knoke & Yang, 2008).

Structural Holes

A structural hole is another concept that is rooted in graph theory and used later in SNA approach. A structure hole within a network happens once a node has connections only to another node belongs to a disconnected component. Where as it has no connections within its component. The result is this node belong to one special node through only one node (Wasserman & Faust, 1994).

3.6.10 Summary of SNA Measures

A brief history of SNA research was discussed earlier in order to get full understanding of it origin. This helps to understand SNA method, measures, and different level of analysis of the social network. Additionally, understanding relationship among individuals and their connections helps in investigating the social structure of the network and the hidden patterns of their behaviors and influences.

The relationships in the networks and the reason for their importance are being studied. Based on the given definitions and explanations, it was established that the use of the SNA concepts is useful to the learning environments and in the study of student's relationships and interactions within the LMS. The most promising concept is that of centrality measures.

Table 3.1 provides a summary of definitions of the main SNA measures (Passmore, 2010).

3.7 Social Network Analysis Software

Several software packages were introduced to study, represent, and analyze social networks. These packages varies from only analyzing social network mathematically

Table 3.1: Summary of most important SNA measures adapted from (Wasserman & Faust, 1994)

Measures	Definition
Density	is the percentage of the connections that are actually found in the network to all ones that can occur.
Centrality	is related to the position of the node regarding to the overall nodes within the network.
Centralization	is related to the extent of which the all the connections are centered around one or a few specific nodes.
Cohesion	is related to the extent of finding subgroups or clusters within the network.

to analyzing both mathematically and visualization through sociograms generation. These package are separate from the method of collecting data for generating social networks. Special packages were developed for analyzing and visualizing larger social network (Huisman, 2003). This section provide a summary of some of the well-known packages for network analysis.

There are many examples of SNA software packages. These packages provide both, mathematical analysis and visualization. Selecting which package to use depends on the nature of the data and purpose of analysis. The following items represent a classification of SNA package application: (1) network visualization through sociogram, (2) network mathematical analysis, through three main level of analysis: node level, subgroup level, and whole network level, and (3) simulation and instant visualization of social relations. Some packages are only built for social network visualization with minimal mathematical analysis capabilities (Huisman, 2003).

Basically, choosing which software to apply is based on the type of the formulated social network and what is needed to be studied. In order to apply analysis and visualization to the network, data first needs to be collected and transformed into an appropriate format to be further analyzed. Second, there is a need to define the level of analysis based on the purpose of analysis. The following subsections illustrate these two ideas. Researchers apply SNA measures to find out the different characteristics of the network and its members. If a network is highly dense or sparse and how subgroups i.e. clusters are found. Also, to understand the overall behavior of network members and how they are connected i.e. network patterns (Bertini, 2005).

3.7.1 Data Visualization

As stated earlier, the main two aspects of SNA are visualization and mathematical analysis. The analysis can be done through the SNA measurements and calculations,

and they were discussed in section 3.6. The following section illustrates how the visualization is done.

Data visualization is defined as the collection of techniques and applications to generate images or graphs that facilitate understanding the hidden relations and patters between social networks. Network visualization is considered as an approach for grouping large amounts of quantitative data and summarize them in one single format. The visualization approaches have many advantages like realizing new hidden relations among network members and finding out new methods for exchanging messages (Freeman, 2004).

Moreover, studying the visual representations of social networks i.e. sociograms and discovering new relations between individuals themselves and between different subgroups and clusters. There are many objectives for studying sociograms, they are summarized as follows:

- discovering subgroups and clusters that are formed based on shared criteria,
- discovering centralized actors with higher influence over the other network members,
- discovering actor that have the role of connecting nodes belong to different clusters and play an important role in keeping the network connected and not breaking down, and
- discovering the isolated individuals with minimal number of connections (Freeman, 2004).

Network visualization is a visual representation for network through nodes and connections between them. Visualization also present additional information on the nodes connections themselves like node shapes, line width or color can convey some properties like connection strength. Some additional information which are result from quantitative analysis can be presented on the final sociogram can facilitate general understanding (Freeman, 2004).

Moreover, SNA software packages can be classified based on their analysis purpose and abilities, first general analysis purpose like UCINET, Pajek, NetMiner, second others having some are built for specific purpose and criteria. The criteria for classifying these package are done based on visualization approaches, mathematical measures offered, and network formation methods (Huisman, 2003).

The following subsections provide a brief description of the highly used SNA software packages. They are UCINET, Pajek, Gephi, and NodeXL. Mainly, all these packages share common characteristics, like extracting data and some visualization approaches. The differences are illustrated within the sections. A summary can be found at the end of the section.

UCINET

UCINET[3] (Borgatti *et al.*, 2010) is one of the earliest SNA packages and widely used for analyzing social networks. It is SNA analytical tool developed by Steve Borgatti, Martin Everett and Lin Freeman (Borgatti *et al.*, 2002). Data manipulation is provided by many format and matrices format in order to generate the matrices from the inserted data. It involves also spreadsheet editor in order to generate sociomatrices to be later analyzed. Regarding mathematical analysis aspects, UCINET have a wide number of SNA measures like centrality measures, clustering and subgroup measures, cohesion measures, and many more. Also, UCINET allow many both ego and whole network analysis methods. UCINET only allow mathematical analysis methods only. However, applying visualization of social networks is done through NetDraw. NetDraw is developed as additional and supplement for UCINET (Borgatti *et al.*, 2002).

NetDraw[4] is a visualizing software for social networks and it works in correlation with UCINET. It accept social networks of different formats like Excel, UCINET, or Pajek files. Moreover, it can visualize and operate larger networks up to 32 767 nodes. Finally, NetDraw provide some analytical measures that can be applied to the visualized graphical social network like centrality measures, clusters detection, and isolated nodes. Both UCINET and NetDraw works properly on Windows operating system and can be adjusted for other operating systems (Borgatti *et al.*, 2010).

Pajek

Pajek[5] is another software package for analyzing large social networks. Pajek allow analyzing social network that are converted into files of NET format. The NET network format is a simple representation of large networks along with the sequence of nodes and all their connections (Batagelj & Mrvar, 1998).

Moreover, Pajek is a free software and is designed by University of Ljubljana and aimed for analyzing large social networks. Pajek accept various input formats for further analysis, like network processed through UCINET format, NET Pajek files, and Pajek projects with extension (.paj) (Batagelj & Mrvar, 1998).

Moreover, Pajek allow social networks analytical approaches both mathematically and visually. For mathematical analysis aspects like centrality, density, clustering coefficient, etc. It has less analytical approaches regarding cliques. Regarding visualization approaches, Pajek provide different graphical configurations approaches (Batagelj & Mrvar, 1998) .

[3] http://www.analytictech.com/
[4] http://www.analytictech.com/netdraw.htm
[5] http://vlado.fmf.uni-lj.si/pub/networks/pajek/

Gephi

Gephi[6] is an open source analytical package for analyzing and visualizing social networks. It is a software for analyzing and handling social data in order to discover social structure of network. Gephi accept wide range of data formats for visualizing, analyzing, handling, and representing social networks. Moreover, Gephi provide various visualization approaches, especially $3D$ visualization and also in real time changes presentation (Bastian *et al.*, 2009).

NodeXL

NodeXL[7] is a free open source analytical software package. It accepts social networks as input that are represented in edge list format where each edge represent a pair of the node and the relation among them. The edge lists can be manipulated through Excel editor format. Moreover, NodeXL provide additional properties for visualization through varying link or relation color, thickness, or label added. Also, changing node properties like shape, size, or color of these nodes. Through that different types of message can be understood only from network sociogram (Dunne, 2011).

Other than various visualization capabilities, NodeXL provide additional analytical measures. It involves different SNA measures like degree centrality i.e. degree, betweenness, closeness, and eigenvector, clustering coefficient. However, NodeXL lacks some advanced SNA measures (Dunne, 2011).

3.7.2 Discussion about SNA software

The most common approaches for studying and analyzing social networks are done through analyzing social networks through matrices, mathematical and statistical methods as long as visual representations of the network. The mathematical and visual representations are an important approach for studying the social structure of social networks (Freeman, 2004).

Studying social structure of social networks involve understanding the way nodes are connected, their relations are formulated, and how the grouping can be extracted. Hence, the main virtue of SNA is understanding how communication and message exchange is done among social network. SNA approach have some main characteristics, like social relation data, visualization of social network, and mathematical analysis of social structure among the individuals (Freeman, 2004).

[6] http://gephi.org/
[7] http://nodexl.codeplex.com/

3.8 Data Mining, Educational Data Mining, and Learning Analytics

As a result of the huge development in Web technologies and applications which is marked by the introduction of Web 2.0, larger amounts of online data can be gathered from various online applications. Moreover, wide use of Web 2.0 applications in education and learning practices (Tim Berners-Lee & Lassila, May 2001). Nowadays, there has a large amounts of online data through many sources like VLE in schools, universities, and learning institutions. All of that arises the need for analytical approaches for mining large databases in order to get valuable information.

Data Mining (DM) is a computing and analytical approach for analyzing databases through discovering and extracting hidden data patterns. DM include a set of well-known methods, such as decision tree construction, rule induction, artificial neural networks, Bayesian learning, logic programming, and statistical algorithms. Other important methods and applications are statistics, visualization, clustering, classification, association rule mining, sequential pattern mining, text mining, etc. (Romero & Ventura, 2007).

Basically, DM is based on exploring hidden relationships patterns from large databases. DM is a branch from Knowledge Data Discovery (KDD). KDD aims in extracting useful and valuable information from large databases. The well-known KDD approaches are gathering, processing preparation, analyzing, and interpretation of the data (Romero *et al.*, 2008).

Specifically, the analysis approach part of the DM is performed through these methods like prediction, clustering, process mining, text mining, and SNA. SNA is an important research approach for understanding relations among any studied case. The main idea of SNA is measuring and investigating relations and roles among individuals or objects that are formulated as networks. SNA define the studied phenomena through representing it as a social network where the entities are represented as nodes and the relations among them as links between nodes. All of the interactions within social networks are studied based on social network theory principles. Different types and kinds of social relations are studied like friendships, marriage, business relations (Romero *et al.*, 2008). This thesis study focuses on using SNA as an analytical tool for analyzing learning networks within learning environments.

Basically, DM was originally applied in analyzing e-commerce and business relations within organizations. It was also used to study the relationships, preferences, and behaviors for customers for any business approach. Recently, it has been applied into e-learning problems and applications. DM applications, methodology, and techniques are applied into learning and education problems which facilitate understanding students and learners behaviors and learning styles and preferences. DM basic aim is uncover hidden relations within the studied problem, this can high a huge impact

on educational aspect. This can be translating in finding new plans for improving learning practices and students outcomes and most importantly enhancing teachers and students communication (Hanna, 2004).

As stated earlier, DM was mainly applied for e-commerce applications and just recently applied into learning and educational problems. The distinction between employing DM in learning practices other than e-commerce or business applications. These distinctions can be seen in the following points:

- Providing advice and comments for students as well as developing new learning plans in the classroom.

- Studying e-commerce applications are aimed in finding new ways in boosting incomes. On the contrary, analyzing learning practices using DM methodology aims in improving learning practices, plans, students' learning outcomes, and improving the methods for in class communication (Barneveld *et al.*, 2012).

Hence, the basic goal of applying DM techniques in educational aspects can be beneficial in handling the increased amounts of online learning data, learners' profile information, and larger number of online learning systems. DM techniques assist in extracting valuable information from huge databases. Another main objective is to understand the communication between teachers and students within course interactions. The output of DM analysis also generates a model that represent the interactive communications between students and teachers. Understanding and analyzing these models can help uncovering relations within learning environments (Castro *et al.*, 2007).

Specifically, applying DM techniques and applications within learning and education aspects gained a lot of attentions. Applying DM within educational aspects have introduced two new analytical approaches, they are Educational Data Mining (EDM) and Learning Analytics (LA). EDM and LA are derived from DM techniques but concentrating on analyzing data about students within courses or learning contexts (Han *et al.*, 2006). The following subsections provide a short overview of EDM and LA.

SNA along with its methodologies and measures are considered one of DM applications. Social networks are extracted from relational database which later represented as graph. Social networks can be used to study problems and be represented through large graphs. Social networks and DM are used to study problems from various research fields like economics, biological networks, and webpages. Furthermore, SNA not only focused on studying the behavior of individuals in the formed networks but also analyzing their interactions. SNA can specifically be used to investigate collaborative learning problems and the interactions within them (Han *et al.*, 2006).

3.8.1 Educational Data Mining

Educational Data Mining (EDM) is is a methodological approach that is derived from DM and it is defined as:

> concerned with developing methods for exploring the unique types of data that come from educational settings, and using those methods to better understand students, and the settings which they learn in[8].

EDM first introduced in 1995 as a result for large amounts of online data that are subject for analysis (Romero & Ventura, 2007). On the contrary, EDM is developed as a new research approach and many conferences and workshops have been held to develop this approach. It is developed for discovering relations and models within learning context and online learning environments. The aim is to get a better understanding of relationships of students and their communications within course (Romero & Ventura, 2010). The results obtained from EDM can help in improving online learning environments, improving relations between students and teachers, and finding new improved learning plans. All of that have a huge reflection in improving learning practices (Zaiane, 2001).

Specifically, EDM methodologies are seen as applications from various research backgrounds such as statistics, data modeling, and data visualization. The basic applications of EDM were built on statistics and visualization. Extracting and discovering models of the analyzed data is considered an important approach for understanding the problem context and relationships among the members (Baker & Yacef, 2009)

As stated previously, EDM applications are derived from DM applications and techniques, however, they are applied into learning environments and context. Its main objectives to uncover relationships between students and teachers and the learning models of how students interact in the learning environments (Baker & Yacef, 2009).

SNA as an application of EDM studies relations between individuals within the network. The use of SNA within learning context is considered as very important advancement in data analysis approaches. These advancement in data analytical approaches is an improvement not only analytical approaches but also social and pedagogical dimensions. This can provide a full understanding of the learning context in all dimensions socially, analytical, and pedagogical (Granovetter, 1973).

Moreover, SNA is examined as one of the applications for EDM. SNA measures and theories can be implemented within learning context as a special case in order to mine and get useful information extracted from learning environments. All of that shape the base for EDM which is used to extract model that represent relationships within learning networks. The main importance of EDM can be seen in many criteria

[8] www.educationaldatamining.org

like analyzing online learning interactions, group formation, communities dynamics within learning environments (Rallo R. & J., 2005).

Furthermore, SNA has an essential role in understanding collaborative and cooperative learning context as well as understanding students, tutor, and teachers parts and roles. Moreover, extracting models that represent patterns of students interaction with themselves, with their teachers, and how they participate in learning activities and learning materials. Applying social network theory into learning context is seen in extracting undiscovered relations in course like basis of communities formation, isolated students with interactions, and factors of motivating student to highly participate in educational activities (De Laat *et al.*, 2007).

Some of the research goals and objectives of EDM is summarized in the following items (Bienkowski *et al.*, 2012):

1. Discovering learning models which represent how students interact, participate, and behave within educational and learning activities within learning environment. This gives an overall understanding of students' behavior, learning styles, and preferences.
2. Examining the results and responses of different learning strategies and plans on students' outcomes and course dynamics.

Nowadays, it is witnessed on the high development on online learning activities and the increased numbers of learning environments along with the use of social media in the learning context. This can be seen as a result in the high advancement in learning technologies and educational programs. The wide spread of online learning applications results in large databases that convey useful information about students such as their personal information, grades, and their participation in the studied learning environments. EDM and SNA can be seen as an essential analytical tool in this era. Specially, analyzing and understanding the interactive models of students behaviors and performance. This can be seen as an early warning system to find out the students with learning problems and poor academic results. All of that allow finding new learning plans to improve learning outcomes (Bienkowski *et al.*, 2012).

3.8.2 Learning Analytics

Learning Analytics (LA) is another concept that is define an analytical approach for aggregating online learning data and analyzing their performance and interactions within online environments. LA is concerned with studying and analyzing students participation in any learning system in an indirect way through gathering the databases of these systems. The objective of this analysis is mainly to find out any association between the students performance in the course or activities and their academic outcomes, discovering models of students interaction and comparing them with previous models of older students. The need is to find out newer plans

for improving students performance. LA focuses on analyzing large databases from learning institutions for extracting useful information, patterns, and models to improve learning within thee universities (Norris, 2010).

Furthermore, LA is rather a new term that describes analytical approaches in educational context. There are many definitions describing this approach, one definition is provided in the LAK11 conference website[9] which states that: "it is the measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs".

Siemens (Siemens, 2010) defines LA as aggregating, manipulating, and analyzing student generated data from any learning online activities for discovering underlying models of their interactions and finally providing recommendations for better results. Another definition of LA is provided by Johnson (Johnson *et al.*, 2011) which states that LA is an approach for understanding large amounts of data aggregated electronically from their interactions in the studied learning environments. Mainly, the aggregation process is done under administrative control and without students' awareness. The analytical process is done under the objective of evaluating the learning process, providing recommendations, and figuring out the problems in the academic process. All the previous definitions of LA share the objective of extracting useful and important information from learning activities in order to understand their behavior and performance. And most importantly providing recommendations to improve the learning process.

All the previous stated definitions of LA vary but all share being an analytical approach in analyzing data in learning context. The objective is extracting models that represent the social interactions of students in any educational context. The resultant learning models allow providing advice and recommendations for improving learning process as well as discovering poor performing students. The results of applying LA approach can be seen in many dimensions, like models of students interactions, analytical information, and visualization of these interactions. These dimensions can be very helpful for learning institution for improving communication between students and teachers and motivating students to participate more in learning activities (Ferguson, 2012).

Basically, LA different approaches are considered largely important in analyzing courses with large number of students. Usually, in these large courses, it is difficult to maintain personal relations with students. It will also be rather difficult to point out students with poor academic results or those who are not willing to participate at all in other activities (Ferguson, 2012).

There are a set of familiar uses of LA approaches especially in identifying students role and predicting their learning outcomes. The output of this analysis can be useful for teachers for identifying difficulties in specific activities, assignments, or

[9] https://tekri.athabascau.ca/analytics/

quizzes. Accordingly, this teachers to introduce new recommendations to apply changes in the plans and activities in order to overcome these difficulties (Norris, 2010).

Basically, SNA and EDM can be considered similar approaches since both focusing on analyzing data focused on extracting and analyzing relational data. SNA is a more general approach through its theories i.e. social theory and can be applied to any problem from different research areas as long as the problem is represented as social network. However, EDM is restricted to learning context and studying learning environments. On the other hand, EDM and LA seem to be similar in their approach and context sometimes they are used interchangeably. EDM are based on DM techniques i.e. classification and clustering but applied for educational context and students problems. Whereas LA are based on both DM techniques and SNA for enabling visualization techniques for representing interactions within online learning environments as graphs. More specifically, LA is used for analyzing data extracted within learning contexts, for example analyzing online courses, PLE, social media for learning context, students portfolios, etc. (Chatti *et al.*, 2012).

According to Downes (Downes, 2010) the older approaches for studying learning environments are always simple like basic statistical measures or number of page hit. However, analytical approaches for learning environments can study and analyze students input, behavior and performance within any learning context. In the study by (Long & Siemens, 2011), they defined some classifications for analysis level for learning problems, such as: using SNA for analyzing course level interactions, EDM used for finding learning model of students interactions for providing recommendations for learning problems. Also, the result of these analysis can be served in proving adaptive learning content according to student's learning styles.

4 Chapter 4 SNA Applied to E-Learning Problems

4.1 Introduction

Learning is a social process which requires interaction and participation of all learning members such as students, teachers, and learning materials. These interactions occur in learning environments; these environment can be designed to support either formal or informal learning approaches. Formal learning approach includes course and classroom-based learning along with lectures, homeworks, quizzes, and communicating with teachers. On the other hand, informal approach includes relations between learners through social media services, sharing the learning materials, working in groups, all done away from grade.

Maintaining social relations is essential for learners' development in their own field. Therefore, social learning theory explains the social relations that arise from interactions within learning environments. One of the earliest social theories is the famous *six degree of separation* which demonstrates that the farness between two persons regarding their personal relation is quite small. Social network theory studies the structure of relations formed from interacting with individuals, and how these structures affect the individuals themselves along with their knowledge capacity (Zhang & Tu, 2006).

SNA, as a research methodology, was highly applied into social and behavioral sciences. Wasserman and Faust (Wasserman & Faust, 1994) credited SNA importance for its features of studying social structures among individuals along with the characteristics of these structures. Social network analysts, therefore, studied different types of social relations that occur among different individual groups not only personal relations but also using resources and exchanging ideas. Social analysts also explored the characteristics of social patterns and these patterns give explanation to any studied phenomena (Haythornthwaite, 2002).

Most recently, SNA theory and measures measures are used to study learning problems and learning environments. SNA measures are used to investigate the interactions

among students themselves and student and their teachers. SNA measures are not restricted to social networks (e.g. Facebook, Twitter) and the resultant social structures. On the contrary, it can study any relation represented as networks. This study investigates learning networks and the ideas of applying SNA measures to learning networks emerged from students' interactions.

Furthermore, many studies have discussed applying SNA into learning context. For example, a study by Haythornthwaite (Haythornthwaite, 1999) used SNA to explore they types of network pattern formed from interactions generated between distance students i.e. within distance online courses. This study concluded that it is important to form study social network from students groups and studying their characteristics. This enhances strong relations among students especially within large networks.

Additionally, SNA is a quantitative method for studying social network patterns. This allow studying the characteristics of these patterns, such as finding out which node (accordingly which person) have a central position in the network. Also, finding out the nodes that play an important role in delivering information to the network members, and the absence of this node cut the information flow within the network (Wasserman & Faust, 1994).

This chapter presents the most recent studies which used SNA as a research method for e-learning problems. Also, it classifies the learning-related-problems to which SNA are applied. This study shows the reasons of the increased attention of social aspects in learning process and the relation to the new approaches of learning.

4.2 Social Relationships within Learning Context

Social learning theory (Vygotsky, 1978; Wenger, 1998) stated that learning best occur through social interactions and communication between learners. This theory gives high priority for maintaining social relations within learning environments and encouraging participation among learning peers. Studying these interactions give an indication of learner's behavior within their own environments either online or courses.

Studying social relations within educational context is stated back to (Vygotsky, 1978) who highlighted that learning occur through participation, sharing, and interacting with others. As a development of (Vygotsky, 1978), the social-constructivism theory was introduced. It allow learners to participate in creating learning material, through the activities and tools within learning environment.

Lately, there has been an increased attention and attention of social aspects in learning. New terminologies have evolved like *social learning* (Brown & Adler, 2008), *social learning networks* (Haythornthwaite, 2002, 2006). These concepts opened up a new

research field that studies another learning aspect through investigating studying learners behavior, performance, and the relation with teachers.

4.3 Present State of using SNA for Learning Problems

This study describes how SNA can be used as an analytical tool to study social relations within learning contexts i.e. either environments or activities. The development in educational context creates new challenges and questions which can be studied using SNA. In the field of e-learning, SNA as a research method, has taken off but mainly in the form of visualization and analysis.

SNA has been widely used in social science research whereas studying learning networks by means of SNA is still in its beginnings (Haythornthwaite, 2011). The following subsections illustrate some studies which applied SNA to education in general and learning environments in specific.

4.3.1 SNA Applied to Online Learning

There have been several studies that used SNA to gain an understanding of classroom activities and online courses. However, most of these studies investigated only interactions within the course not the behavior of learners themselves. For example, a study by Nurmela (Nurmela *et al.*, 1999) extracted log files generated from an online course and the formed networked accordingly. This study focused on centrality measures, degree and betweenness. Degree measure results provided an indication of how active the student was in the course, i.e. higher number of messages sent and received. Students with higher betweenness measure value where the ones who controlled information/news distribution among the whole class.

Another study by Martínez (Martínez *et al.*, 2002) studied the importance of online tools as a communication mean between teachers and students from one side, and among students themselves from the other side. It was done through analyzing the event logs generated from web conferencing session among one class. However, Hiltz (Hiltz & Turoff, 2002) suggested that this communication pattern is not ideal for online student learning.

A study by Reffay (Reffay & Chanier, 2002) argued the importance of Distance Learning Management System (DLMS) as an approach to minimize teacher's role within distance online course. This study showed that the significance of studying students groups and the characteristics of these groups. The authors studied the groups through defining clusters and cliques. Thus, studying these cliques giving an understanding of the relations within the course as a whole not only individuals separately. This helps to discover any problems within the course.

Analyzing data extracted from LMS become one of the hot topics lately that studied relationships taking place in learning environments. Many studies took place in this field and investigated students behavior. Thus, they can provide insights for students' results within the course. One study conducted by Campbell (Campbell *et al.*, 2007) found a strong correlation between student presence in online activities with their outcomes in a designed model of their performance.

Another study by Dawson (Dawson, 2006) studied the data extracted from online environment and specially discussion form messages. This studied the association between number of message exchanged and the tendency of building communities within the course. Also, Weaver (Weaver & Morrison, 2008) examined a great correlation between student participation level within the course discussion forum and their grades.

4.3.2 Student Performance

Many aspects of student's behavior within online course were studied from SNA point of view, for example, understanding different communication styles in collaborative environments. A study by (Cho *et al.*, 2007) studied students' communication styles within learning environment. They analyzed learning communities by form them in social networks and later analyzing networks using. In their findings, they found that there is a strong relation between students position in the social network and their final grades.

Moreover, a study by Martínez (Martínez *et al.*, 2003) used two analysis approaches SNA and quantitative analysis methods to study interactions in classroom. They used two data sources the traditional questionnaires as long as data log files of the online interactions. SNA measures facilitates the understanding of the students willingness to participate and interact within the class.

In another study(Yang & Tang, 2004), the author studied the group structures and patterns formed from students actions within an Information System (IS). SNA was used as an analytical tool to analyze group cohesion of the groups. The results of studying group cohesion gave an indication of how strongly connected the students within the formed groups. The students who are members in the same group share similar learning characteristics.

There are quite a lot of researches on virtual communities and some of them investigate learning virtual community. A study by Hamulic (Hamulic & Bijedic, 2009) determined characteristics of the virtual social network for overall communication for information technology students. They found out that some of the most successful students are the stars of the three analyzed networks.

Furthermore, Newman (Newman, 2004) analyzed different types of collaborative networks of different research fields physics, biomedical research, and computer science. The formulated networks represented the relation between scientists in these

fields. This study investigated different characteristics of the networks like, who is the scientist with the central position within the network? who is the one with the largest numbers of written papers?

4.3.3 Building Learning Communities

The previous section 4.3.2, presented some studies that focus on finding a connection between learners' participation within online activities and their academical performance. Moreover, students activities within online environments forms learning communities. These online groups or communities were studied using SNA measures. Many questions raised from studying the learning groups or communities like if there is a correlation between communities characteristics and the success in online courses as well as investigating the characteristics of the members of these communities.

For example, a study by Dawson (Dawson, 2008) studied SNA learning communities through SNA centrality measures. It found that applied to learning networks social network found a connection between degree centrality measure and community formation.

Similarly, Rovai (Rovai, 2002) investigated the importance of studying learning communities in order to understand how learning is taking place. Higher percentage of interactions can be seen as important factor of finding learning communities within courses. These interactions are results of talking among students and studying them provide an indication of the success of online course.

A study conducted by Cho (Cho *et al.*, 2002) studied SNA measure of prestige applied to study learning environments. Applying prestige measure give a clue of how fast information and/or news can be distributed among the environment. The obtained resulted showed the importance to motivate student participation for more effective environment dynamics.

Another example is the study done by (Aviv *et al.*, 2003) which examined Asynchronous Learning Networks (ALNs) within specific course. Their findings showed that students have an essential part in initiating the conversations within the course. On the other hand, the teacher have minimal part throughout conversations. Hence, studying the characteristics of these networks characteristics of these networks like cohesion enhances knowledge building and sharing process.

4.3.4 Frequency of Information Distribution

In order to gain an insight of students' style and behavior within courses, it is vital to study the structural patterns of their interactions within any online activities

within learning environment. These online activities varies depending on the plan of the online course.

Many studies (De Laat *et al.*, 2007; Erlin *et al.*, 2009; Willging, 2005) considered using SNA as an analytical tool for learning environments and for detecting students' involvement in online courses. These studies attempted to apply SNA measures analyze the network structure emerged within online activities, i.e. discussion forums. Applying SNA measures can give another dimension in understanding course dynamics. Course dynamics can be studied based on different levels either studying the importance of students position within the network, or studying the characteristics of the overall learning communities. Thus, the result of these measures can give a view of how students interact and participate within these online activities.

The study by Haythornthwaite (Haythornthwaite, 1999) considered the studies in the same direction in applying SNA measures in learning environments. This study investigated specifically the case of distance learning course. In its findings, this study showed the importance of the relations maintained within smaller groups i.e. networks, as a part of the whole course. These smaller groups maintain stronger connections between their individuals which means the individuals within this group can get the information faster than others, as well as they exchange information between each other not the other groups.

4.3.5 Group Cohesion

The study by Reyes (Reyes & Tchounikine, 2005) examined two SNA measures, centrality and group cohesion, to study social cooperation and communication within a learning context. They used centrality and cohesion to measure and understand the characteristics of the extracted forums' networks. Students who got higher centrality results where the ones controlling the discussions and replying the most number of messages. Having a few number of dominant within the discussions may affect frequency of finding communities. This have an impact on the rate of data transmission among the participants.

Consequently, Reffay's study (Reffay & Chanier, 2003) explored other aspects of analyzed network. Specifically, studying the network based on the individual's level such as discovering the inactive students and the features of group members. This study points out the importance of group structure within the course. The finding out the of such as the such as found out that the isolated people, active sub-groups, various roles of members in the group of communication structure, both individual level and group level to assist the tutor in following collaboration with group.

Accordingly, the study by Saltz (Saltz *et al.*, 2004) explored SNA measures to study class connections. This study focused on degree centrality measure for understanding

students' connections within the class. They showed a relation between online participation and their achievements in the course.

As previously stated, SNA allow analyzing the network based on two levels: among the individual level which studied the behavior of the individuals based on their position within the network. The second level is based on the overall characteristics of the network and the extent to which individuals are clustered into subgroups. Studying subgroups formed from the students within learning environment helps to get an overall insight of the dynamics of the class. Diverse types of relations can be discovered from investigating learning communities from online learning activities. For example, exploring how individuals are connected can provide answers about how fast news are dismissed within the class. Also, how individuals share and collaborate within and work jointly to solve an exercise. All of that enhance information distribution base within the class (Daniel *et al.*, 2008).

Moreover, Shen (Shen *et al.*, 2008) considered the factors that influence online learning interactions. This study explained a high importance of studying class relations in maintaining a sense of community and joint group relations among students within the class. Also, this study utilized Rovai's Classroom Community Scale (CCS) and SNA measures (density and network centralization) to understand community patterns of learning environments.

4.4 Summary

Recently, SNA is applied into learning environments and problems to study the social aspects learning environments. Applying SNA measures can be summarized into these following aspects (Sie *et al.*, 2012):

1. Visualizing and representing the connections and relations among students as a social network. A social network of the interactions where the nodes represent either students or teachers. The visualization of the network can give information and understanding of how the structure and patterns of the class activities.

2. Analyzing the extracted social networks through using SNA measures. These measures allow mathematical analysis of the network in which the individuals with the important characteristics can be spotted, i.e. the ones with largest number of interactions.

3. Studying these networks can help understanding what are the aspects that affect students collaboration within the class.

4. Finally, analyzing the social networks mathematically and study the emerged patterns can allow teachers to find new plans and approaches for teaching in the class. Thus, finding new approaches for motivating students for learning.

Consequently, the base of SNA measures and theories are largely compatible with learning environments and theories, especially, social constructivism which is the basis of some use LMS (Sie *et al.*, 2012). These assumptions can be summarized as follows (Knoke & Yang, 2008): (1) analyzing and understanding social connections gives a better understanding of the whole network rather than focusing on personal characteristics of each individual, (2) visualizing the networks, how the networks are connected, and patterns of network can provide full understanding of the dynamics and relations among network, and (3) analyzing and visualizing social networks are a dynamic relations and they are variable throughout time.

5 Chapter 5 Methodology and Design Patterns

This chapter discusses the methodology, data sources for analysis, research design approaches, and the theoretical framework in this study. It clarifies the data mining approach used to extract and later on analyze the studied data.

The focus of this research is to study networks extracted and formulated from students' interactions within some of Moodle's modules, i.e. discussion forums and wikis. These networks are analyzed using SNA measures. Specifically, SNA measures allowed quantitative analytical method, such as the analysis of different levels of connections between students themselves and students and teachers within courses. The importance of SNA resides in understanding the social structural patterns of students' interactions within the LMS courses (Sie *et al.*, 2012).

Most universities use LMS to facilitate and support classroom teaching. LMSs have common services for managing and saving students and teachers information and learning materials. These information form large amounts of data that are stored in relational databases within universities' systems. Utilizing these databases can help gaining information about students and their connections within LMSs(Black *et al.*, 2008).

The objective of this study is to demonstrate the applicability of SNA to evaluate the social roles in online learning activities. The results of SNA measures point out various characteristics of students i.e. most or less active student and their roles within the course. Besides, these results provide an insight of how the connections among the course members formulated. This research presents a case study on investigating interactions between the course participants that are taking place in two online modules in LMS i.e. discussion forums and wikis.

5.1 Introduction

SNA is a research method used mainly in social sciences, but in recent years it gained importance in other fields. Using SNA as a research methodology gives a complete understanding of the dynamics within the course and connection patterns of the extracted social network. This research studies the participation structure and level in online learning modules (Haythornthwaite, 2005).

Besides being used in different social science research areas, SNA has recently been applied to LMSs and learning environments. This approach gained a lot of attention due to the large amounts of online generated data along with the development in Web 2.0 online services and applications (Safran *et al.*, 2007; Ebner, 2007).

In the early days of online learning, researchers were focused on creating and distributing course materials along with creating computerized learning programs. However, as a result of the technological advancements the need arose to analyze students' interactions, connections, and learning styles and how all of that affect their achievements.

The choice of analyzing discussion forums and wiki activities was because of their collaborative and interactive type. These activities are applications of social constructivism theory (Kop, 2011). Thus, networks are formulated based on the data extracted from the module interactions, specifically from forums and wikis. Subsequently, these networks are analyzed to discover many indicators for understanding course dynamics. For example, one of the indicators used is the frequency of the interactions taking place within the studied modules.

Consequently, discussion forums and wikis networks are represented in sociograms i.e. a graph. The sociograms illustrate the structural patterns of discussion forums and wikis social networks. These sociograms are also subject for mathematical analysis using SNA measures like density, centrality, and cohesive subgroups among students.

The purpose of the visual representations i.e. sociograms of the emerging networks is to help teachers understand students' online engagement in a course. That can enable teachers to change the teaching plans for the course based on the studied sociograms (Sie *et al.*, 2012). The individuals within sociograms represent all participants within the course students, teachers, and tutors participating in the online discussion forums.

SNA is a quantitative method of analyzing the social relations of online users. Its application in education has created a new perspective for educational research, and gradually revealed a lot of information on the educational phenomenon, helping us to understand some of the problems students might have (Sie *et al.*, 2012).

The following sections represent the methodology along with data collection and analysis approaches.

5.2 Research Methodology

This research is taking place on the basis of Moodle, evaluating the social interactions within the LMS developed by the TU Berlin (called ISIS) based on the learning networks formed from students connections within ISIS modules.

The research approach of this study is a quantitative. Generally, research methods can be classified into three main categories:

(1) Quantitative Research,
(2) qualitative Research, and
(3) mixed-Method Approach (Creswell, 2002).

Quantitative research is an approach suitable for evaluating social experiences through out mathematical measures and techniques. The quantitative methods consist of surveys and experiments. Qualitative research on the other hand provides a full understanding of the studied case and how individuals interact. Qualitative research method are based on questionnaires and interviews. Finally, the mixed-method approach explores an intermediate approach between quantitative and qualitative approaches (Amaratunga *et al.*, 2002).

The choice of research method should be made in relation to the research objectives. The most used qualitative research methods are presented by Hunter (Hunter, 2004):

1. Action Research: studies the effect of change on individuals or groups,
2. Case study: studies a phenomenon according to all the elements where it took place,
3. Ethnography: studies different obtained notes of a problem throughout time,
4. Grounded theory: studies a specific problem without considering the previous conditions,
5. Narrative inquiry: studies all the cases that are connected and common to individuals under study.

This study is based on qualitative research techniques, and more specifically, a case study approach. This approach is applied because it is appropriate for interpreting, investigating, discovering relations within a specific context. Mills (Mills *et al.*, 2010) stated that using this approach provide an understanding of specific teaching plans and how the plans have an impact on the outcomes. In other words, studying the effect of new parameters by applying them on a specific case and later on can be generalized into the whole environment.

5.2.1 Case Study

Case study research approach is a method that explores a The case study research method is an empirical method that investigates a new emerged case or problem within any research approach, throughout testing different hypothesis to get a full understanding of the whole case (Yin, 1984). Moreover, this approach enables researchers to study specific assumptions and their effect on the case while stabilizing all the others. Studying specific assumptions within case study provide a complete view of how the interactions within the real environment will take place (Dalcher, 2004).

In a case study research, there are six steps that must be utilized apply this research (Soy, 1997):

1. defining the research questions,
2. selecting the phenomena to be studied along with appropriate data aggregation and analysis methods.
3. gathering and preprocessing data to be studied and later on analyzed,
4. explaining the obtained results and finding the connection real life case.

Therefore, this research is based on case study approach, as it studies the interactions between the participants of some of the modules within Moodle (i.e. wikis and discussion forums) in courses from the perspective of social network analysis.

5.3 Data Collection Methodology

In literature, several approaches were introduced for collecting data for analysis, using social networks measures. These approaches vary from the traditional ones like surveys and questionnaires to the automated ones, such as log files which are extracted from either websites or databases. The traditional data collection approaches are very time consuming tasks, and not always accurate. They are even more complicated when dealing with large amounts of data. Interviews or surveys for example are greatly time consuming and difficult to be employed in many different circumstances (Gretzel, 2001; Romero *et al.*, 2008).

As a result of the large development in Internet communication technologies i.e. Web 2.0 and social media, computerized approaches for gathering online data become highly used. For example, automated log files holds many useful information to understand the online system. A system logs are files that consists of data about system users and their interactions within the system i.e. number of times of access, access time, and the accessed resources (Romero *et al.*, 2008).

The most suitable method to collect data for analysis in a learning context is the learning systems themselves. These systems capture a wide range of students'

information, from personal information, to their interactions within the course, as well as their data such as grades, forums posts, and wikis entries.

The case study presented in this study focuses on analyzing data produced from a LMS activities, specifically ISIS (Moodle at TU Berlin. This study represent users within a ISIS course as nodes within a social network, and the connection between two nodes represents the relation established between them.

5.3.1 Moodle's Database and Data Collection Approaches

This study is taking place within the LMS at TU Berlin, a Moodle is called ISIS. Moodle is a good example for studying activities taking place within learning environments. Moodle keeps a detailed logs of all activities that students participate in. Moodle usually stores its log data in a relational database. LMS log files are the most promising source of automatically gathered online learning data (William, 2006).

These logs consist of information of all events performed by Moodle's users, such as communicating in forums and chats, reading study materials or blogs, taking tests or quizzes etc. The users of this LMS are students, tutors, teachers, and administrators which interact among each other and with the learning materials under course boundaries (Romero *et al.*, 2008). The objective of this study is to investigate students' activities in ISIS and to discover the hidden social network created from groups of students with similar patterns of behavior.

Moodle save every action the participants did such as accessing the system, download material, or even sending email in the system log files. These files are categorized by course, participant, day, and activity. Database administrated and teachers of specific courses can access these information to refer back to students' activities like who has been active in the course, what they did, and when they did it.
Furthermore, Moodle classify log files according to their type as well such as the number of posts in discussion forums, number of messages exchanged along with the ones who read/wrote them. Participants' information are stored in the database as well, for example participants' message within forums, chat, or wiki pages (De Laat *et al.*, 2007; Willging, 2005; Erlin *et al.*, 2009; Laghos & Zaphiris, 2006).

5.3.2 Early Data Manipulation

Data used in this study is extracted from the database of the Moodle LMS system used within TU Berlin, ISIS. Confidentiality is the obstacle to access data from ISIS. This is due to the fact that these are personal and private data belongs to students as their grades and names. Consequently, all the names were anonymized and all connections allowing the reconstruct of any personalized information about the students were removed. To further ensure that privacy will be respected a

document was signed, which stated that the data will only be used for research and analysis purposes, without the intention to recreate any names or information about students.

5.3.3 Information Extraction and Analysis

Basically, the analyzed data were mined and extracted from ISIS database for different courses under investigation. The data belongs courses of different semesters. The processing approaches are based on data mining methods for extracting or dealing with large amounts of data. The extracted data belong to participants' communications from two main ISIS modules i.e. discussion forums and wikis.

The required data extracted from ISIS is stored in a relational SQL database. SQL queries are used to extract the required data for further analysis. A MySQL Query Browser was used for extracting the needed data from ISIS. It allows the creation of tables, editing, and inserting data either by viewing or writing SQL statements.

5.4 Moodle Database Structure

All user interactions that took place within discussion forums and wikis were stored in ISIS's relational database. As stated earlier, the data were retrieved from ISIS and later preprocessed for further analysis. In order to retrieve information from large databases, data mining approaches are needed.

Basically, SNA analytical method was used in this study as an application of data mining approach. This is done using a specific SNA software. This study uses UCINET software to analyze and process the extracted data, and NetDraw to visualize the networks and create sociograms (Borgatti *et al.*, 2002).

Moreover, UCINET is used to convert the results out of SQL queries into a matrices i.e. sociomatric form which is the suitable form to represent the extracted networks. The sociomatrices represent the relations by creating an actor-by-actor matrix. Accordingly, these matrices are subject for analysis by applying SNA measures. Sociomatrices were imported to UCINET (Borgatti *et al.*, 2002) for analyzing the data. Another software offered with UCINET is NetDraw. NetDraw is a software for visualizing social networks as graphs where nodes represent actors, and lines represent relations among them (Hanneman & Riddle, 2005).

Moodle Architecture

As previously, Moodle stored detailed information about each event done by its participants, such as: the accessed module, the type of action (read, write, or modify), the date and time, and the IP address of the computer. Moreover, Moodle operates on web server which are based on PHP programming language and database systems like Apache, PHP, and MySQL (William, 2006).

Moodle does not store log information about their participants and modules in a relational database, MySQL is used by Moodle. Database systems are the best systems for saving and later on retrieving large amounts of information. Most importantly, these systems are good for minimal error rates in retrieving appropriate information (William, 2006).

The following figure 5.1 represents the layers of which Moodle consists. The mostly used ones are: discussion forums, chats, wikis, or messages. These modules are subject for analysis due to their collaborative type. On the contrary, quizzes, assignments, and feedback could also be considered, but it is still difficult to define interaction models on the students' activities in them. Since there interactions are based on one way communication, initiated by the teacher to the whole class students.

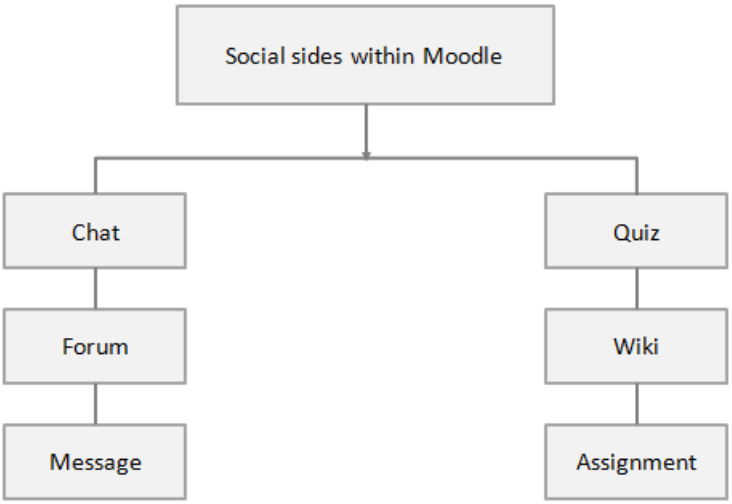


Figure 5.1: Moodle's Modules and Services

5.5 Methodology Steps

The methodology used throughout this study is based mainly on DM approach, specifically educational data mining, since the aim is to extract data about students interactions within LMS. Specifically, the studied and analyzed modules are discussion

forums and wikis within ISIS. This study is built upon the data mining steps presented by (Romero *et al.*, 2008), and in this study are adapted for being used for learning environments. These steps are:

1. Collecting data from ISIS: Moodle stores extensive log files in a MySQL relational database.

2. Pre-processing of the collected data: transforming the collected data into an appropriate format in order to be processed by SNA software. The data contain users information they have been anonymized to eliminate any possibility to get their personal information.

3. After collecting and modifying the required data from ISIS, the last step is the analysis step to understand the hidden information from the extracted information, in this study is applying SNA measures to understand the formulated networks. In this step we used the software UCINET to create and visualize the SNA results.

4. Finally explaining and analyzing the results got from the analyzed networks. This allow obtaining an overall overview of the dynamics within the studied courses.

Moreover, the main steps that were followed throughout this study can be seen as: First, collecting data from ISIS relational database about the two modules under study: discussion forums and wikis. Second, applying appropriate SQL queries in order to get the needed information. Third, applying their results as an input into UCINET for further analysis. The output from processing data by this software took two forms: the analytical part and the visualization one through sociograms.

For each forum and wiki extracted modules, directed patterns of relationships were extracted and later on represented into an adjacency matrix. For both modules, the relation is defined as who replied to whom in the forum, and who changed or edited any of the wiki pages. The course sociomatrix i.e. adjacency matrix contains information about the students' mutual relations based on their actions on either discussion forums or wikis.

5.5.1 Networks' Analysis Approach

Data is extracted from the databases using SQL queries and then processed into matrices to later be analyzed and interpreted. The main approach in this study is based on analyzing the data collected from ISIS's records. The measures used in this study include centrality, cohesion, clustering, and cliques of students' networks from the analyzed courses.

5.5.2 Adjacency Matrix Definitions

The network is formulated as an adjacency matrix A where $A_{ij} = 1$ if there is a relation between two nodes i and j, and otherwise $A_{ij} = 0$. The generated adjacency matrices are symmetric such that $A_{ij} = A_{ji}$ (Passmore, 2010; Sie *et al.*, 2012).

In order to visualize discussion forums and wikis networks, their relations must first be of interactions among participants, their interactions are mapped into a two-dimensional matrix. A matrix of a network of size n is a square matrix such as $n \times n$ (Passmore, 2010). The network is formed as a graph with nodes representing participants in the learning environment (e.g. students, teachers, and tutors) and edges representing the relationships among them (within discussion forums and wikis).
The relation between nodes A and B is marked by 1 in the cell (A, B) and (B, A) if there is a relation between them, and 0 otherwise. If the direction in the relation is important, then there is an edge from A to B. On the contrary, if direction does not matter, there is only a line from A to B (Passmore, 2010).

5.5.3 Analysis Part using SNA

This research studies learning networks formulated from two learning modules within LMS: discussion forums and wikis. These learning networks are extracted from the students interactions and entries within these modules where the nodes represent participants (e.g. students, tutors, teachers) and the link between any two nodes represent the existence of a relation among them.

Specifically, this research studies these learning networks extracted from courses offered within LMS. This research is based on case study approach for investigating these learning networks. The LMS defined for TU Berlin university ISIS has been considered as a case study in this thesis. The information is extracted from discussion forums and wikis modules within different modules. These courses were chosen as being the ones with the maximum number of interactions taking place. The information extracted through SQL queries from ISIS relational database.

This study uses SNA software UCINET, for data analysis and manipulation. UCINET works in combination with NetDraw visualization software. The main function of UCINET is to prepare and manipulate matrices which represent the studied learning network. As stated in UCINET, this software can also be used for analyzing and manipulating: one-mode networks, and two-mode network (Borgatti *et al.*, 2010). Basically, all the learning networks defined for this study, discussion forum and wikis learning networks are one-mode networks.

6 Chapter 6
Networks and Results

In this chapter the results are obtained from analyzing the networks generated from student's interactions within two interactive modules in Moodle (ISIS); discussion forums and wikis. This chapter includes definitions of forums and wikis, as well as the measurements applied on the basis of the networks. The network extracted from ISIS is used for further analysis.

The focus of this study is to consider how SNA measures can be used to examine students' and teachers' interaction patterns and styles within the learning modules in LMS. A network is formed from the dynamics of interactions between students and teachers from discussion forums and wikis which are learning modules in ISIS. Collection and analysis of the data are done through both mathematical and visual approaches.

These basic questions will be answered: are students interactions' within the modules related to their learning and academic outcomes? How is a node's position in the network related to the student's role within the class i.e. central student? Generating sociograms of the learning networks will reveal hidden interactions between nodes i.e. students, teachers, or tutors. Through this study, various quantitative methods were used, such as: centrality, cohesion, clustering, and overall connectedness of the networks. These measures studies the learning networks from different level of analysis, such as individual level i.e. centrality measures, group levels i.e. cliques and clustering, and the whole network level i.e. centralization.

One of the objectives of this study is to visualize the interactions among the studied networks, and hence provide another approach to the evaluation of students' interactions within the class. It will show that the students who are more active within the class occupy a central position within the network, and if those who do not attend classes, assumed to be isolated and having few or no connections, are not participating in the forums and wikis.

Most recently, SNA as an analytical approach in educational context and analyzing learning problems. Learning problems are studied based on social network theory in which any problem is represented as a network where nodes represent students and links represent connections between students themselves and teachers of the course.

There are many different types of social relations that can be extracted from learning environments. SNA can be used for analyzing these social learning problems based on many analytical level such as individuals roles, subgroups existence, overall network cohesion, characteristics of social structure, and many more (Reyes & Tchounikine, 2005).

Some factors that affect the results, the structure, and the pattern formation of the networks (forums and wikis), are the grades of the students, the existence or not of teachers, and the limitation of the size of the network. The results can also be explained by considering the number of assignments submitted and their grades by students.

Looking thoroughly at these interactions and how active students are within these modules can give an overview of the course dynamics, more specifically point out the less active students and those who are not participating in the course modules. It can also distinguish between high and low performing students (teachers can use this knowledge to adapt their methods) or discover tendencies in the creation of learning groups or communities. It is found that less active students were the ones who did not register for the final exams, and accordingly their marks are low.

6.1 Introduction

SNA was carried out to study and analyze the social relations among the participants within LMS (students and teachers) as well as studying the levels of participation in the ISIS modules, i.e. discussion forums and wikis. Chapter 3 provides an overall introduction to SNA and its applications. Two main approaches were studied: (1) data extracted from ISIS's modules; discussion forum and wiki in specific, and (2) analysis and visualization of sociograms discovered from the social networks of discussion forums and wikis.

Furthermore, throughout this study the three main levels of analysis for studying the social networks extracted from learning modules of ISIS, they are:

1. analyzing the whole network characteristics i.e. size, network centralization, and density,
2. analyzing the subgroup characteristics i.e. clusters, subgroups, or communities,
3. analyzing individual's characteristics i.e. centrality measures of degree, betweenness, and closeness.

Both discussion forums and wikis are examples of collaborative learning tools are the basis for group work among students. They are considered as an application of the social learning theory among representing the interactive parts, in this case students and teachers, as node of the networks and the relation among them as links. This

facilitate understanding the overall relations among the learning process (Kop, 2011; William, 2006).

The following subsections are organized as follows: general statistical information on Moodle database, network definitions, and finally the discussion and analysis of the results.

6.1.1 General Information from ISIS

In general, Moodle offers many learning modules to be used in learning environments: quizzes, discussion forums, wikis, assignments, instant messages, and chat. However, not all of them are highly used by teachers and students much. This is due to the facts that teachers still use Moodle as space to distribute their learning materials and students tend to interact more through informal spaces (i.e. blogs, Facebook).

LMSs mainly support the creation and distribution of learning materials to students, the preparation of quizzes and assignments, the beginning of discussions, or organize distance courses. These modules generate a large amount of data, mainly stored in relational databases. Such data provides insights into student's behavior and learning styles, and also serve as a review for teachers and tutors. Other common data sources, such as personal students' and teachers' profiles, results and user interaction data, are also available in the system database (Romero *et al.*, 2008).

Thus, the LMS stores large amounts of data in a relational database, which can be challenging and rather difficult to get concrete information about students. Moodle offers a statistical module, but this is relatively weak. This study uses SNA as an analytical tool for studying dynamics in courses in learning environments. The results of studying the course dynamics from SNA perspective can provide teachers with valuable information for changing their plans and strategies in the classroom and increase productivity (Romero *et al.*, 2008).

One of the challenges faced is that wikis and forums are used rarely, and it was difficult to find courses that used them both. Wikis are new in Moodle, ISIS in specific, and are only used by a few courses. Moreover, courses that use wikis mostly are targeted to teachers or tutors as a part to motivate them to largely use Moodle's modules within their courses. Those were the courses with higher interaction values within wikis. The most important challenge however is that Moodle is still considered purely as a repository to download learning material i.e presentations and it is hard to start conversation between students.

6.1.2 Networks Definition

Basically, through this study two types of networks are defined; wikis and forums. These two specific modules were chosen based on the their interactive and collaborative nature. These networks are extracted from specific courses.

The criteria for selecting the considered courses are:

1. those with largest number of students,
2. those with forums and wikis with the largest number of interactions, participation levels, and students using them,
3. those that use assignments and quizzes modules at the same time.

As stated earlier, the extracted data from ISIS is transformed into an adjacency matrix. Later, the generated network is formed into a directed networks of the nodes (students, tutor, and teachers) and the relations between them. These networks, either forums or wikis network, serve as an input parameter into UCINET and the NetDraw network analysis software.

6.2 Forums' Networks Analysis and Discussion

A discussion forum is another model of Computer Mediated Communication (CMC) which represent a type of asynchronous online interactions between individuals i.e. students, teachers electronically through any defined online network. Discussion forums enable students, teachers, and tutors to send and receive messages under specific discussion topics i.e. problem solving, programming problems, or news distribution. This communication model is considered as example of collaborative learning approaches (Cole & Foster, 2007).

For each course two forums were analyzed in order to find out on which there were more interactions i.e. higher participation levels. These forums varied from being question/answer forums to homework submission ones. Accordingly, some of these forums have students as the central player in the network, and others have teachers as the central ones.

Through the following subsections, the computed measures for forums' networks are described. Overall network cohesion, degree of nodes within the network, and subgroups characteristics of networks are all computed for the defined network.

Forum's Network

A discussion forum is an activity module in Moodle (ISIS) where students and teachers can exchange messages/information/messages. Every single post along with

who posted it, time of creation, time modified, and message content, is saved in the ISIS database. After forums' information have been extracted through SQL queries, it is inserted into an adjacency matrix that can be further analyzed through UCINET software. Similarly, the generated matrices are imported into NetDraw software to generate the required network sociograms.

The defined forum network is based on the relation between two nodes, i.e. students, teachers, or tutors, regarding their communications. This relation is based on who started any conversion, and who sent/receive a message to whom within the threads in the forum.

6.2.1 Forums' Networks Information

11 courses were analyzed to extract networks from discussion forums. For each course two forum networks were extracted. Originally, the plan was to analyze, and later on compare, four forums for each course. Unfortunately, there are few interactions taking place within the courses, and the result is thus a smaller matrix based on fewer networks.

In order to analyze a forum's network, basic statistics are extracted first, which help to understand the whole basic idea behind the studied courses. Table 6.1 lists the studied courses, the number of students in each course and the semester in which the course was offered.

Table 6.1: Courses used for analyzing forum's network

Course	Course Name	Students
C-1	Algorithmische und Funktionale Lösung diskreter Probleme [WS 10/11]	231
C-2	Blue Engineering	87
C-3	Methodische und Praktische Grundlagen der Informatik 2 [SS 12]	270
C-4	Algorithmische und Funktionale Lösung diskreter Probleme [WS 10/11]	231
C-5	Semantik und Kalküle	27
C-6	Einführung in die Informatik 1 [WS 12/13]	1176
C-7	HTA Online 2012/2	20
C-8	Agententechnologien in der Forschung [SS 12]	40
C-9	MPGI 5: Datenbanksysteme[SS 12]	300
C-10	Berechenbarkeit und Komplexität [SS 2011]	202
C-11	Berechenbarkeit und Komplexität [SS 2009]	116

Table 6.2 shows the number of forums defined for each course. Course C-6 has the highest number of forums, with 19, while C-11 only has 2. There is no direct

relation between the number of defined forums and the size of the later extracted network.

Table 6.3 shows the total number of forum posts. The forums chosen for analysis were the ones with the largest number of posts. Likewise, there is no direct relation between the size of the network and the number of posts considered.

Table 6.2: Number of forums defined within courses

Course	No. of Forums
C-1	11
C-2	10
C-3	5
C-4	3
C-5	3
C-6	19
C-7	15
C-8	11
C-9	3
C-10	2
C-11	3

Table 6.3: Number of posts within forums in the analyzed courses

Courses	Forum-1	Forum-2	Forum-3
C-1	112	60	38
C-2	6330	557	85
C-3	12	45	14
C-4	1178	305	39
C-5	178	132	37
C-6	338	293	274
C-7	93	68	68
C-8	47	18	0
C-9	62	52	23
C-10	84	26	0
C-11	46	35	0

A social network is characterized according to some basic statistics such as its size. The network size is defined as the number of nodes forming the network. The number of relations that are forming the network is not related to the actual possible number of nodes that could actually be formed. The largest number of relations that could be possibly generated in a network of N nodes differs between directed and indirected relation. This is calculated for directed network based on this equation $N \cdot (N - 1)$ (Hanneman & Riddle, 2005).

Table 6.4 shows the sizes of the forums' networks. The size of the network is a factor for the structure of social relations within this network. This table shows relatively small networks, the sizes range from 7 to 38 nodes for the first forum to 0 to 16 for the second one.

Table 6.4: Network size of forums' network along with the number of pairs formed

	Forum-1		**Forum-2**	
Courses	Size	Pairs	Size	Pairs
C-1	17	272	10	90
C-2	18	306	12	132
C-3	44	1892	30	870
C-4	39	1482	2	2
C-5	8	56	2	2
C-6	7	42	2	2
C-7	6	30	4	12
C-8	7	42	6	30
C-9	23	506	4	12
C-10	22	462	2	2
C-11	10	90	6	30

6.2.2 Cohesion

Density is social network measure that computes the level of coherence between nodes within the network. This measure gives an indication of how fast the information/news are distributed among nodes in network. The result of applying this measure represent the level of connectedness between individuals in network who are connected to others but not counting the number of nodes that actually form the network (Hanneman & Riddle, 2005).

Generally, density is a quantitative representation of the overall distribution of the network; it ranges from 0.0 to 1.0. The higher values (i.e. higher than 0.5) indicate that the communication within the network is constant and the speed of information distribution among network members is high.

Table 6.5 shows mean density values calculated for forum networks for each course. The means of the density values were small, and vary a lot for each course. The range of mean values of density for the first forum ranges from from 0.02 to 0.22. For the second network, the values range from 0.003 to 1. Courses C-1 and C-3 have the highest density value for the first forum. Whereas C-10 shows the lowest values for the first forum.

For the second forum, C-5 shows the largest value, with 1.0 (it is a very small network with only two actors), whereas C-9 shows the lowest value of 0.033.

Table 6.5: Density measure values for forums networks

Courses	**Forum-1**	**Forum-2**
	Density	Density
C-1	0.26	0.04
C-2	0.14	0.10
C-3	0.02	0.02
C-4	0.22	0
C-5	0.13	1.0
C-6	0.02	0
C-7	0.17	0.17
C-8	0.05	0.07
C-9	0.03	0.0033
C-10	0.03	0
C-11	0.07	0.2

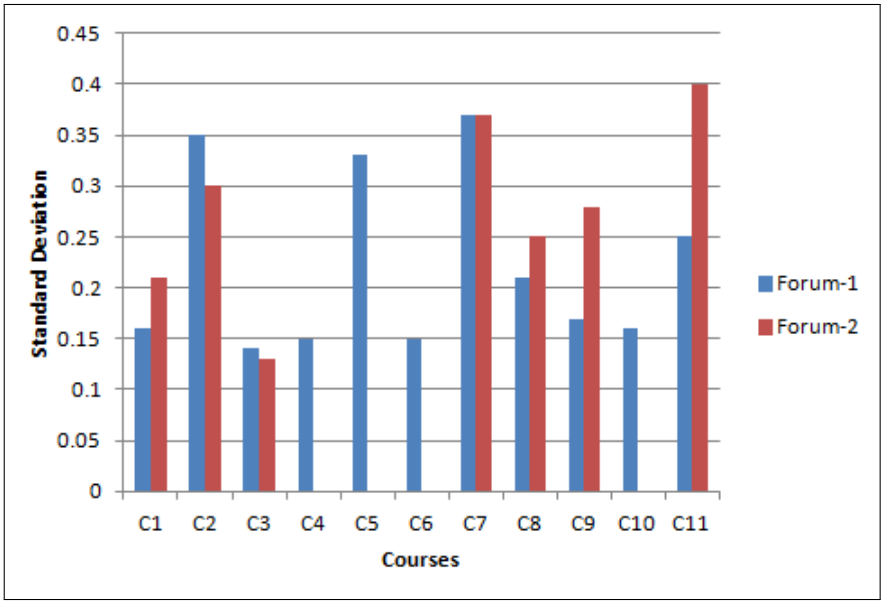


Figure 6.1: Standard deviation values of density measure for forums' networks

Figure 6.1 shows the standard deviation values computed for the density values. The diagram shows a large variance between the values. The largest variance is for the second network of course C-11.

Generally the overall values are very small and, consequently so is the ratio of interactions. Lower density values indicate a low connectivity between participants, while higher ones indicate higher connectivity, more students are participating in this forum. The justification for these small density values is the existence of a large number of disconnected components, less messages exchanged, or a small number of

different users.

The number of students registered within the courses are large, but unfortunately there are only a few participants in the modules. Students are not willing to interact with each other or to start a dialog. One possible reason is that Moodle, ISIS in specific, is still seen as a CMS instead of a communication or discussion system. Students tend to use the Internet to solve some problems, or to interact on a face-to-face basis with their fellows and thus show less tendency to formulate online social relations with class fellows. This is considered as the social constraint for forming the network.

Another important observation, the networks (forums) with higher densities are the ones in which teachers/tutors highly participate in. This emphasizes the idea that forums are more like an informative place to distribute the class news i.e. exam dates and due dates of homeworks, rather than a place for students to discuss curriculum problems and solve the homework.

6.2.3 Centrality Measures

There are three different types of centrality measures, they are degree, betweenness, and closeness centrality. These measures analyze the network based on node level which means they study the characteristics and roles of individuals not and how these characteristics affect the overall network performance. Centrality measure studied the different roles each node have according to the position of the node in the network. The node's position within the network have a connection to many aspects like its importance, power, and influential in the network (Hanneman & Riddle, 2005).

6.2.3.1 Degree Centrality

Degree centrality computes and identifies the node's importance and influence within the network. Degree centrality measures is divided into two types based on the type relations of the network. For directed networks, there are two degree centrality types in-degree and out-degree. In-degree centrality represents the incoming connections and measures the node's importance or popularity in the network. On the other hand, out-degree centrality represents the outgoing connections and it measures the node's influence or power in the network (Hanneman & Riddle, 2005).

Consequently, the results of degree centrality describes the coherence of nodes in the network whether they are homogeneous or heterogeneous. The results of degree measures provide a clue of the variability of degree whether they are low or high (Hanneman & Riddle, 2005). The in-degree of student's online learning interaction is the total number of replies a student obtained from other students, while the out-degree is the total number answers a student posted.

Classifying students according to centrality values gives an indication of how active students are within the course, and points out the less active ones i.e. peripheral or isolated students. The range of in-degree values is larger than that of out-degree values and the standard deviation values of in- and out-degree values are very different.

Table 6.6 shows mean values for in-degree and out-degree values for the forum's network for each course. The distribution of mean degree values for the network of the first forums range between 0% and 16.667% for both in-degree and out-degree. The values for the second forums range from 0 and 13.3.

Table 6.6: Mean values of degree centrality for forums' networks

	Forum-1		**Forum-2**	
Courses	Out-degree	In-degree	Out-degree	In-degree
C-1	0.0%	0.0%	0.4%	0.4%
C-2	2.0%	2.0%	1.08%	1.08%
C-3	2.0%	2.0%	0.533%	0.53%
C-4	2.0%	2.0%	0.0%	0.0%
C-5	13.0%	13.0%	50.0%	50.0%
C-6	2.38%	2.38%	0.0%	0.0%
C-7	16.67%	16.67%	16.67%	16.67%
C-8	4.76%	4.76%	13.3%	13.3%
C-9	4.76%	4.76%	13.3%	13.3%
C-10	2.60%	2.60%	0.0%	0.0%
C-11	6.67%	6.67%	1.0%	1.0%

The mean value of degree for course C-1 is 0% and 0.4% for the second forum network. These forums are called "discussion information" forums. Unfortunately, the participation was extremely low, and the actor with the highest degree value was the tutor responsible for distributing information about the course to the students.

On the other hand, the highest mean value for both forums' networks is for course C-7, the mean value is 16.67. The participation is relatively high in this forum because it is a new one. Here the actor with the highest degree is the teacher. This forum is called an "asking question" forum, where students can post questions for the teacher. However, the participation from students is relatively low.

Course C-10 has the highest in-degree and out-degree values in both forums. This highest value is for teachers in both cases. Generally, the teachers' degree values were always higher, except for the homework submitting forums. The forums where the students have the higher degree values are those with higher grades in the assignments. The range of the grades are between 0 and 10. The students with

higher degrees share the same range of grades, between 6 and 8. Those with low degree values or zeros either did not submit the assignments or got grades below 5.

Figure 6.2 shows the standard deviation of the degree values, and the high variations among them. It also shows that more messages are sent than received. The highest out-degree value is for Course C-5, a forum where students can ask teachers questions.

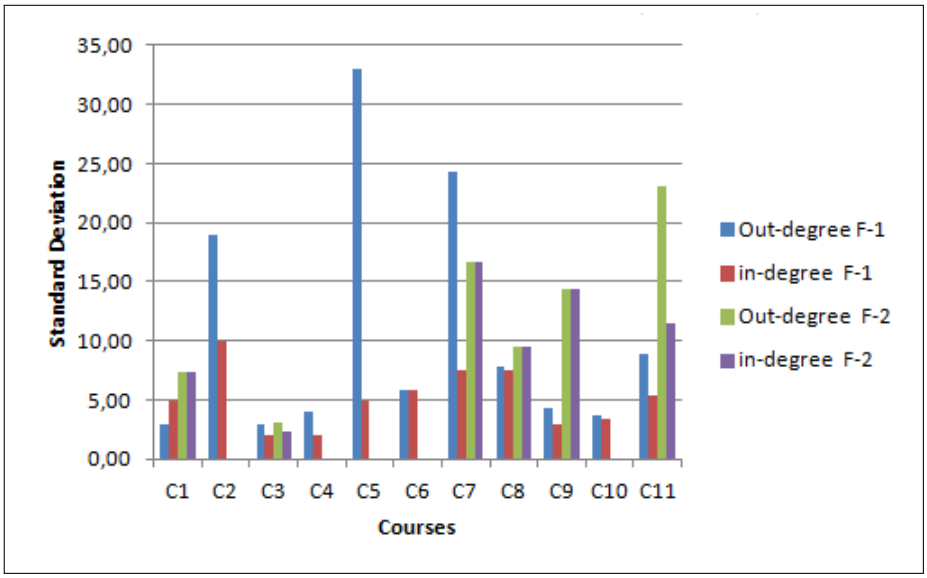


Figure 6.2: Standard deviation values of degree measure for forums' networks

Table 6.7 shows the minimum and the maximum number of in-degree and out-degree values for the forums' networks. The minimum values are zeros for both degree values. The maximum number of received message, i.e. in-degrees is 11 and 6 for sent messages, i.e. out-degree for the first forum. For the second forum, the maximum number of received message in-degree is 7 and 4 for out-degree .

Figure 6.3 shows the network of course C-2 as an example. This figure shows the most central node in regard to degree centrality. The size of the node is related to the values; the higher the degree value the bigger the size. Nodes on the edges of the sociogram represent the isolated nodes, the students who did not participate in the forums, and did not send or receive messages.

6.2.3.2 Betweenness

Betweenness centrality computes the relative importance of the node as a result of its position in the shortest path between any two communicating nodes. The betweenness values for a node describes its high connectivity to other nodes within the network regardless of type of connections being directed or indirected. The

Table 6.7: Minimum and maximum values of degree centrality for forums networks

Courses	Forum-1				Forum-2			
	In-degree		Out-degree		In-degree		Out-degree	
	Min	Max	Min	Max	Min	Max	Min	Max
C-1	0	1	0	3	0	2	0	2
C-2	0	11	0	6	0	4	0	3
C-3	0	4	0	3	0	7	0	4
C-4	0	6	0	4	0	0	0	0
C-5	0	1	0	1	0	7	0	1
C-6	0	1	0	1	0	0	0	0
C-7	0	3	0	1	0	1	0	1
C-8	0	1	0	1	0	1	0	1
C-9	0	1	0	1	0	4	0	2
C-10	0	3	0	3	0	0	0	0
C-11	0	2	0	1	0	3	0	2

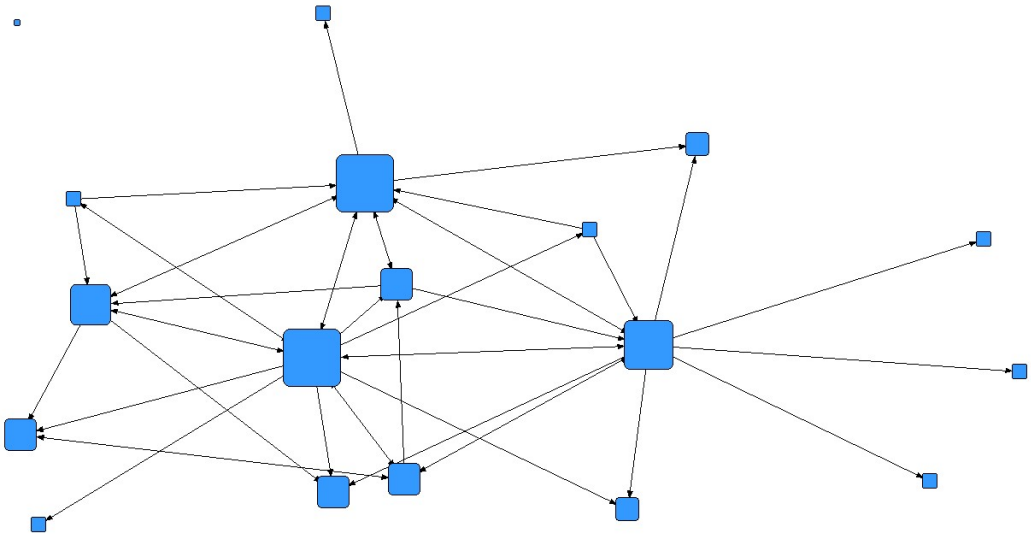


Figure 6.3: Course C-2 of forum-1 network arranged based on the in-degree values

results of this measure points out the nodes that play as broker within the network. The broker is the node on which all the others in the network depend in order to distribute the news to the rest of the network. Moreover, betweenness centrality describes the general connectivity of the network since it computes connectivity of each node to the others of the network. Usually, networks with low density values and larger number of disconnected components have very low betweenness values and sometimes zero (Hanneman & Riddle, 2005).

Table 6.8 shows the mean values of betweenness centrality of the forums' networks. The mean values are relatively small overall, they range from 0% to 15%. This is

due to the low density values among the networks themselves. For courses C-1, C-5, C-6, and C-10, the betweenness values are 0%. Other courses have different values for their first and second forums, as for example C-8, where the betweenness value for Forum-1 is 3.33% and 0% for Forum-2.

Table 6.8: Betweenness measures values for forums' networks

Courses	Forum-1	Forum-2
C-1	0.0%	0.0%
C-2	13.43%	6.12%
C-3	0.11%	0.59%
C-4	0.85%	0.0%
C-5	0.0%	0.0%
C-6	0.0%	0.0%
C-7	10.0%	0.0%
C-8	3.33%	0.0%
C-9	0.22%	0.0%
C-10	0.0%	0.0%
C-11	0.0%	15.0%

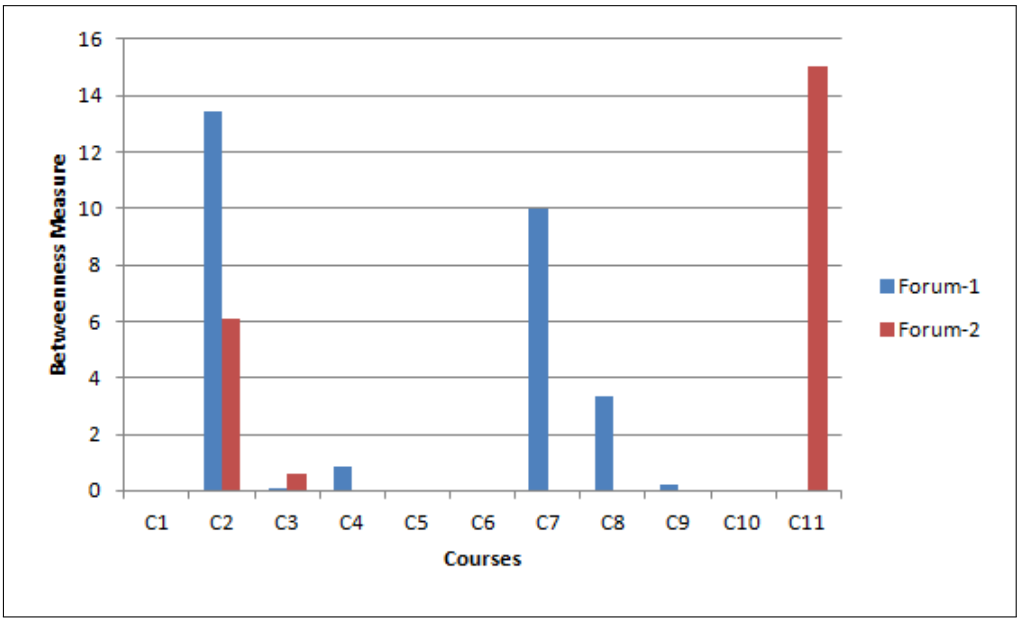


Figure 6.4: Betweenness values for forums' networks

In general, low betweenness values mean that there are no gatekeeper in the network, no specific node were all the communications must go through in order to succeed. The network is thus not in danger of breaking down in the case of the removal of these main nodes. The nodes with betweenness values of 0 indicate that these nodes

are on the edge of the network. The more one actor controls others, the higher the betweenness value. The nodes with higher betweenness values will play the role of bridges for other members in the community/group where the communication occurs.

The overall betweenness values are very low and less informative. This can be related to the low densities (low connectivity among the nodes) of the forum, to the many disconnected components and to the fact that the forums serve the purpose of information rather than discussion.

6.2.3.3 Closeness

This type of centrality is related to the distance measure where it computes the distance of one specific node to the rest of nodes within the network. It points out the importance of the node in the network as a whole or in small regions. For example, a node may be central but only in small region where the node has many connections but only in smaller region. Closeness centrality is also computed according to the direction of relation in the network. For directed network there are in-closeness and out-closeness measures. Moreover, closeness can be calculated for the whole network unless the network does not contain disconnected components (Hanneman & Riddle, 2005).

Generally, another aspect of the network can be understood for applying closeness network. For example, high closeness value is an indication that the node is highly independent in the network i.e. means depending less on others to receive information. Whereas, low closeness value is an indication of node's dependency on others to get any information. Therefore, the standard deviation for the closeness values are taken into account in order to avoid extreme values (Müller-prothmann, 2007).

Table 6.9 shows the mean closeness values of the forums' network. The out-closeness values for the first forum ranges from 2 to 27.18 for courses C-3 and C-7 respectively. For the second forum, they range from 0 (for courses: C-4, C-5, C-6, C-10) to 29.17 for course C-7.

Figure 6.5 shows the standard deviation of the degree values. There is a large variance among the out-closeness standard deviation values. They range from 0 to 20.02 for Forum-1 networks and from 0 to 19.567 for Forum-2. The in-closeness values range from 0 to 0.167.

Table 6.10 shows the minimum and maximum values for in-closeness and out-closeness values for the two forums of each course under study. The maximum value is 71.43 for out-closeness for the first forum in course C-7, as well as for the second forum of course C-11.

Table 6.9: Closeness measures values for forums networks

	Forum-1		Forum-2	
Courses	In-closeness	Out-closeness	In-closeness	Out-closeness
C-1	6.0	6.0	10.47	10.47
C-2	10.0	21.0	9.89	10.57
C-3	2.0	2.0	3.44	3.44
C-4	3.0	3.0	0.0	0.0
C-5	14.0	23.0	0.0	0.0
C-6	14.63	14.63	0.0	0.0
C-7	20.71	27.18	29.17	29.17
C-8	15.35	15.35	17.778	17.78
C-9	4.49	4.49	27.08	27.08
C-10	4.67	4.67	0.0	0.0
C-11	10.67	10.72	22.48	28.37

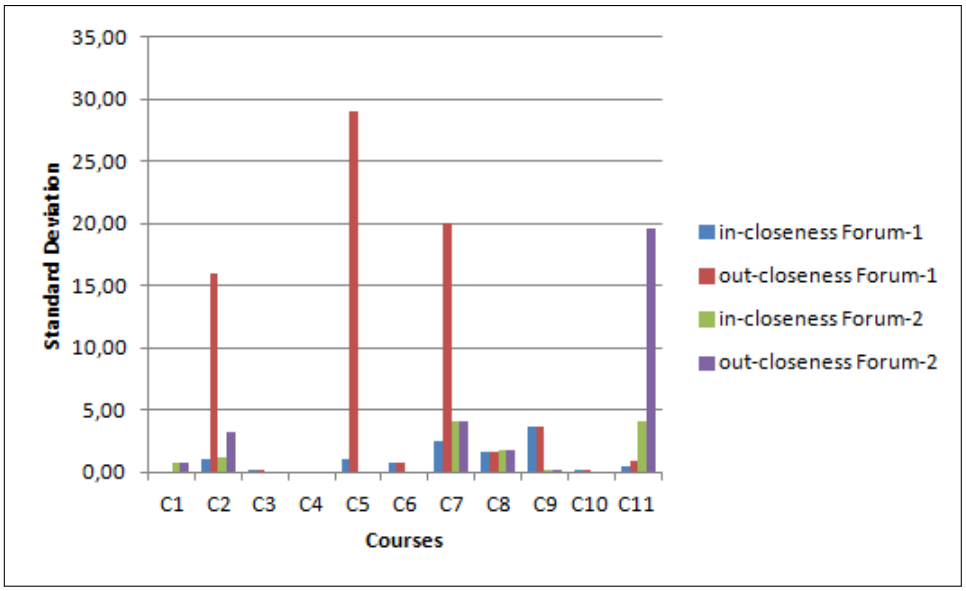


Figure 6.5: Standard deviation values of closeness measure for forums' network

6.2.4 Centralization

Network centralization computes the overall centrality of the whole network not for each individual node. In other words, the result of this measure gives an indication if the interactions within this network are centralized around one or more specific node. Once the centralization values are high, this gives an indication that there is only one specific node that controls the information distribution among all the nodes. The removal of this specific node breaks the network and stopped the information distribution (Müller-prothmann, 2007).

The very centralized networks are those with higher centrality values. The result of

Table 6.10: Minimum and maximum values of closeness centrality for forums networks

	Forum-1				Forum-2			
Courses	In-closeness		Out-closeness		In-closeness		Out-closeness	
	Min	Max	Min	Max	Min	Max	Min	Max
C-1	6	7	6	6	10	12	10	12
C-2	6	11	6	44	8.33	12.36	8.33	16.17
C-3	3.33	3.83	3.33	3.98	2	3	2	3
C-4	3	3	3	4		0	0	0
C-5	0	0	0	0	13	14	13	100
C-6	14.28	16.66	14.28	16.66	0	0	0	0
C-7	16.67	23.81	16.67	71.43	25	33.33	25	33.33
C-8	16.67	20	16.67	20	14.28	19.35	14.28	19.35
C-9	25.0	33.33	25.0	33.33	4.38	4.76	4.38	5.26
C-10	4.55	5.26	4.55	5.26	0	0	0	0
C-11	10.0	11.11	10.0	12.5	16.67	29.41	16.67	71.43

this measure figure out the nodes i.e. individuals and the relation to their position within the network. This is referred to the most important important node i.e. individual within the network. Calculating the centrality measure for the whole network expose valuable information about the network structure and how nodes are represented. A network with high centrality values show that there one or few nodes are controlling the network dynamics, in other words, the network is focused around these smaller node sets. If these few nodes were removed from the network, this made the network to be broken or disconnected. The centralized nodes are the ones who have responsibility of keeping the network connected. On the other hand, networks with lower centralization values have no single nodes that are controlling the connectivity within the network. There are no single nodes which their removal from the network can break it.

Figure 6.6 shows the distribution of centralization values. The central actors were mostly the teachers. There are large differences in in-degree and out-degree centralization measures, especially in course C-5.

When a network is centralized equally, its value is 100%. This may be due to the fact that there are no significant interactions taking place within the forum, all the posts are only from one actor posting new messages to the students in the course. The networks are centralized around tutors whom are responsible either to dismiss the information or the answer questions within the course.

6.2.5 Clustering and Subgroups

The second aspect of studying the network is finding smaller groups aggregated from nodes within network. There are many terms describing subgroups within

Table 6.11: Network centralization values for forum networks for out-degree and in-degree measures

	Forum-1		Forum-2	
Courses	Out-degree	In-degree	Out-degree	In-degree
C-1	4.17%	18.33%	19.75%	19.75%
C-2	56.62%	23.53%	28.97%	19.01%
C-3	14.95%	7.64%	12.37%	8.79%
C-4	14.29%	8.73%	0.0%	0.0%
C-5	116.67%	2.38%	100.0%	100.0%
C-6	16.67%	16.67%	0.0%	0.0%
C-7	52.0%	4.0%	22.22%	22.22%
C-8	13.89%	13.89%	8.0%	8.0%
C-9	15.99%	6.41%	33.33%	33.33%
C-10	12.25%	12.25%	0.0%	0.0%
C-11	17.28%	4.94%	48.0%	24.0%

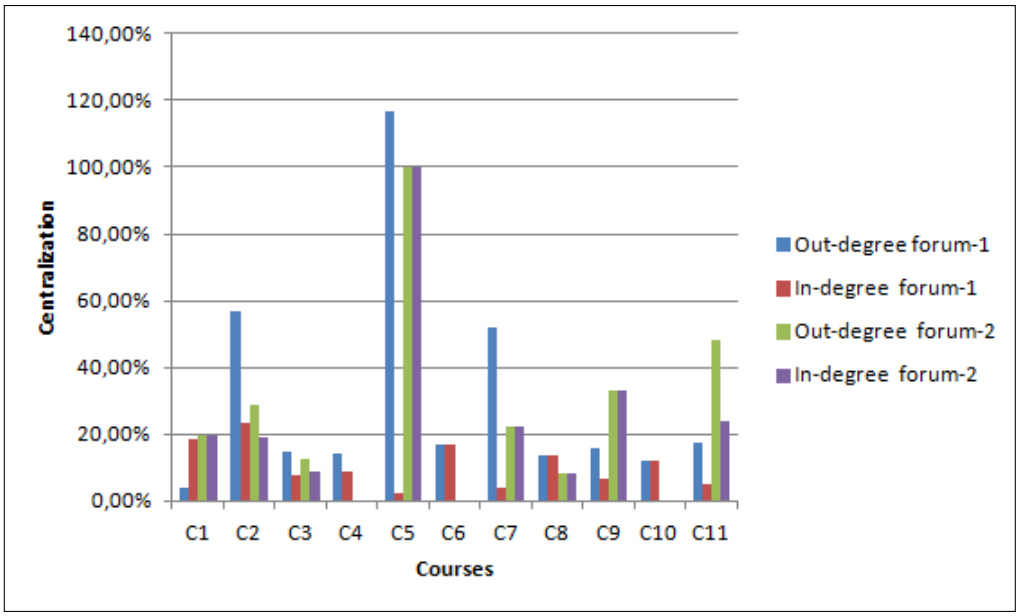


Figure 6.6: Centralization values for forums' network

network like clusters, components, and communities. They are interchangeably and they describe the tendency of individuals to work collaboratively. They focus on finding subgroups within the network: clustering coefficients, components, cliques, and N-cliques (Hanneman & Riddle, 2005).

More specifically, studying and working in groups is the main feature of learning. Students tend to study collaboratively with their peers and then finding out groups within the networks. This can be reflected to learning approaches through the learning theories of social constructivism, connectivism, and social learning. All

of these theories investigate the approach of working within groups along with the characteristics of students belonging within these group. Also, the benefit students and teacher can gain from working jointly on problems (Kop, 2011; Dunaway, 2011).

Basically, there are two main approaches for finding subgroups within any social network, either by detecting communities (i.e. bottom-up approaches) or detecting components (i.e. top-down approaches) within network. The bottom-up approach is defined based on this approach where the process firstly defined any two connected nodes in a small N-clique and then aggregated other nodes afterward until another dense neighborhood, but separated from the reset of the network, is found. The top-down approach on the contrary, is based on studying and analyzing the network as a whole i.e. one big component and afterward detecting the disconnected components and isolated ones (Hanneman & Riddle, 2005).

The following subsections show the result of applying these measures to forums' networks. It also shows how these measures explain the students' interactions in the learning environment.

6.2.5.1 Clustering Coefficient

This type of measure is that computes the tendency of finding groups, clusters, or communities within network. There are two methods for measuring clustering coefficient; either it is calculated for each node of the network, or it is calculated for the whole network as one component. Clustering coefficient for a single node is defined by the density of this node within its immediate region or subgroup region. On the other hand, the clustering coefficient for the whole network is computed as the average of the result of clustering coefficient for each node. The importance of this measure is seen as the degree of finding communities or clusters within the network. A higher values represents a higher tendency of finding groups and higher possibilities that nodes have larger number of connections with others. Moreover, clustering coefficient is an implementation of one of social network theories which is small world problem. Small world problem theory detects and represent the small group or immediate region around one specific node, and it can find and define the local neighborhood within the whole network community (Hanneman & Riddle, 2005).

After calculating this measure for forums' network, it was found that most of the values were zero except in C-2 course for the first forum. The clustering coefficient for this network was 0.622. This trend is explained through the fact that none of the networks are highly dense, they are all composed of many separate, disconnected sub-groups.

6.2.5.2 Cliques

Cliques is another term used to describe the existence of subgroups or clusters within the network. A clique is a measure to define and detect subgroups in a network. It is also defined as the largest complete subgraph within the whole network. The set of the nodes composing the clique have stronger connections with each other comparing to the rest of the network (Knoke & Yang, 2008). Analyzing and detecting cliques can be explained in real life problems as the willingness of individuals to collaborate in groups of other individuals who share similar interests and characteristics.

Moreover, defining and analyzing cliques within networks are related to measuring how fast the information can be distributed among network members. This also depends on the origin of which node initiate this new information. The information is spread faster within highly dense clusters. Usually, any new information is introduced into clusters through an outsider node. The nodes within each cliques are referred to as strongly connected with each other, while the relation between this clique and any outsider node is referred to as weekly connected (Hanneman & Riddle, 2005).

For the first forum, cliques were only found in C-2 (four of them). Therefore, the remaining courses were analyzed to find out subgroups through analyzing N-cliques measures along with varying the value of N. Figure 6.7 represents the clique data set of course C-2. In the network the square represents the clique set, and the circles represent the members of each one of the cliques. The nodes on the far left represent the ones who are isolated and do not belong to any one of the cliques found.

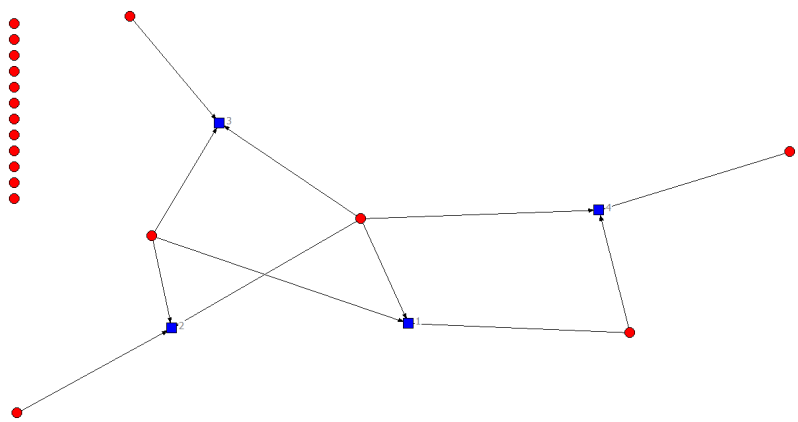


Figure 6.7: Clique sets of course C-2 for the first forum

6.2.5.3 N-Clique

N-cliques is another definition term describing the process of finding clusters or subgroups. N-cliques is simpler and looser definition of cliques such as there is a path length of N or less that is connecting the clique members (Knoke & Yang, 2008). Similarly to cliques it is an indication of how students are willing to work collaboratively in smaller groups.

This aspect is calculated using 2, 3, 5 for N. It was difficult to find larger numbers of N to be studied. The range of the groups was small and this is reflected in the low connectivity (low density values). Table 6.12 shows the number of N-cliques found in the forums' networks.

Table 6.12: Number of N-cliques formed within forums' network - different N values

Courses	**Forum-1**			**Forum-2**		
	2	**3**	**5**	**2**	**3**	**5**
C-1	1	1	1	0	0	0
C-2	4	1	1	0	0	0
C-3	0	0	0	0	0	0
C-4	13	9	5	0	0	0
C-5	1	1	1	0	0	0
C-6	1	1	1	0	0	0
C-7	2	1	1	6	1	1
C-8	1	1	1	0	0	0
C-9	6	5	3	0	0	1
C-10	3	2	2	2	2	2
C-11	2	2	2	2	1	1

The number of these sub-groups is relatively small. C-4 has most sub-groups within its network. There are 13 dyads, 9 triads, and 5 sub-groups in which N is 5. These values are for both forums. In course C-9 for example, it was found that these subgroups were found in its first forum network. The numbers of dyads, triads, and subgroups of 3 are 6, 5, and 3 respectively. It is noticed that the sub-groups contains the teachers and/or tutors within them. The teacher was the one who replied the messages within this network.

Figure 6.8 is an illustration of a clique tree of two cliques participation for course C-9 for the first forum. The ID's of the students who participated within them were omitted due to privacy issues. This diagram re-orders the actors within the network in a way that the ones who belong to the same clique are placed close to each other.

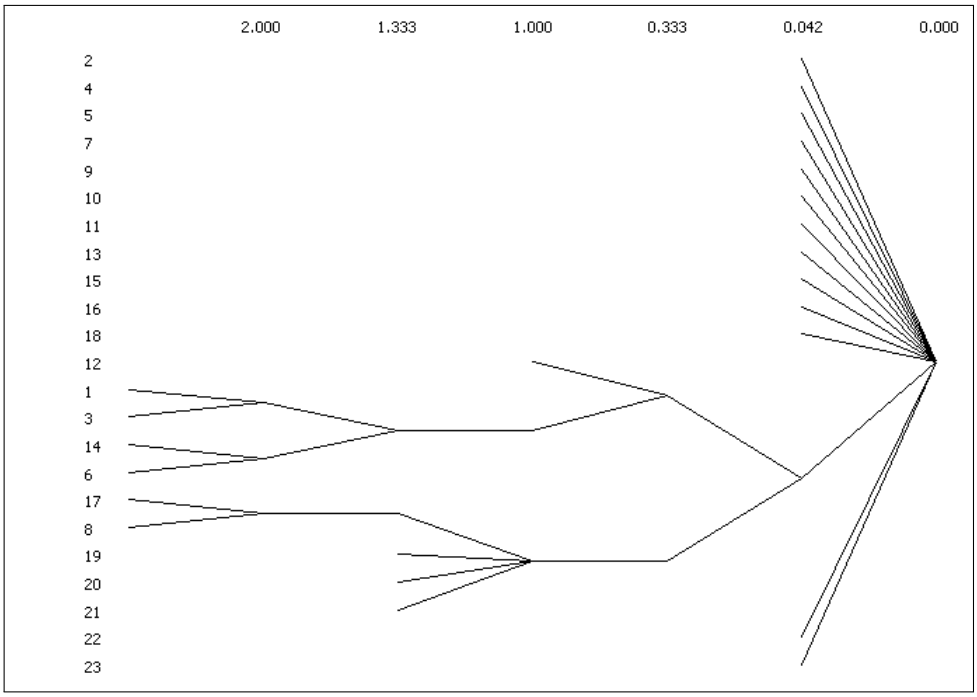


Figure 6.8: Tree diagram for N-clique cliques where N is 3 for course C-9 for the first forum

6.2.5.4 Components

A component is basically a cluster i.e. cluster of the network in which all of its nodes have links to each other by at least one connection regardless of being direct or indirect. Based on this previous definition the node without any direct or indirect connection to at least one node is referred to as isolated i.e. separate component. The larger number of components found, the higher indication of nodes to be aggregated in components. This measure is essential for detecting the overall connectivity of the network. Defining components facilitate detecting cut-points components. Nodes that are marked as cut-point in the network are the ones their removal breaks the network. There are another term related to component, it is disconnected component which represent a component that have no connection of any type with the rest of nodes of the network. Finally, the specific node that connect two or more separate components together are referred to as bridges (Hanneman & Riddle, 2005).

Moreover, a component is the basic representation of community or subgroup where its nodes have connections with each others but not with the rest of the network. The components can be classified according to the type of the links forming them either direct or indirect links. Components that are formed based only on the directed links are called strong components. Whereas the components that are formed regardless of their links directions are called weak component (Hanneman & Riddle, 2005).

All the previously defined concepts and measures are used to analyze and define

ego-networks. Finding out the different roles each node occupy within the network facilitate understanding their behavior and how they act. For example, finding broker nodes represent weak components with weak connections. Additionally, finding nodes that are considered bridge, isolates, or structural holes within the network. Each one of these roles have a definition related to the real life problem that is studied according to social network theory (Hanneman & Riddle, 2005).

Table 6.13 shows the number of weak and strong components that were found in the forums' networks. Course C-3 has the largest number of components, both strong and weak, for both forum networks. There are 44 and 29 strong components for the first and second forums' networks respectively. Generally, the networks of the second forums are very sparse and no components can be found. The higher the number of components, the smaller the number of the members in them.

Table 6.13: Number of strong and weak components computed for forums' network

	Forum-1		**Forum-2**	
Courses	Weak Comp.	Strong Comp.	Weak Comp.	Strong Comp.
C-1	10	17	0	0
C-2	2	10	0	0
C-3	8	44	15	29
C-4	8	39	0	0
C-5	1	8	0	0
C-6	6	7	0	0
C-7	1	6	0	0
C-8	5	7	4	6
C-9	8	23	3	4
C-10	10	22	0	0
C-11	4	10	0	0

6.2.5.5 Ego Network Measures

Ego network are network extracted from the whole one and is constructed from one single node together with its relations. This specific node is called ego and its extracted network is generated from the viewpoint of this node, this is why the ego network is also called personal network. The personal network is made up from the ego and its alters which refer to the nodes that are directly connected to the ego. The ego network is seen as the network structure from the point of view of the ego (Hanneman & Riddle, 2005).

Table 6.14 shows the measures that are calculated for in ego networks for the two basic networks. The basic measures are:

1. size of ego network which is the number of its alters,

2. summation of directed ties,
3. the density of this specific ego network i.e. personal network,
4. diameter of the personal network which is the longest distance to the ego,
5. summation of weak components and their relation among the network,
6. summation of brokers,
7. betweenness centrality measured for this personal network (Hanneman & Riddle, 2005).

This measure is calculated and displayed only for one specific network of a specific course. One example for this measure is the ego network of the teacher within the network of course C-3, for the first forum. The study concentrates on the ego network with the highest number of actors. Course C-2 for example has two main egos within its network, which are the teacher or the administrator.

Table 6.14 shows the calculations that were done for one specific ego-network (the teacher being the ego). The density measure can also be calculated for the ego networks. The higher the values, the denser the network. As density decreases, more structural holes are likely to open in the social fabric.

Ego	Size	Pairs	Density	Broker
Tutor	5	20	0	10
Student-1	4	12	0	6
Student-2	1	2	0	0

(a) Course C-3 for first forum

Ego	Size	Pairs	Density	Brokers
Student	4	12	0	6
Student	2	2	0	1
Student	1	1	0	0

(b) course C-9 for first forum

Table 6.14: Ego-network measures for courses C-3 and C-9 of forums' networks

For course C-3 three different ego-networks were calculated. The first for the tutor, the second for a student with a high grade (student-1), and the last for the one with lowest grade (student-2). For course C-9 however the egos are from the students with the highest grade values, student-1 and student-2, while student-3 is the one with the lowest grade.

6.2.6 Network Visualizations

Studying social networks is done jointly both mathematical and visual approaches. The mathematical analysis is done through the different measures and their results was previously presented.

Moreover, studying social networks visually is done according on graph theory by representing the interactions taking place through graphs. The individuals i.e. nodes together with their connections i.e. relationships are represented in sociograms. Sociograms describe the social structure of the networks. The importance of network visualization resides in understanding the social structure of the network and real world problem is represented. Studying social network through visualization is considered as a supplement and addition to the mathematical analysis (Niu, 2010).

There are many observations resulting from the examination of the different sociograms of forums' networks. For example, students formed small discussion subgroups, whose subjects were mostly personal opinions and examples. In some of those moments they replicated the dominant pattern (star pattern) while in other moments they replicated the line pattern. In the first pattern one student is the center of the debate, while in the second pattern students were essentially talking to just one or two colleagues. The students at the end of the lines were just interacting with one other colleague.

Figure 6.9 shows the forum networks of course C-3. This figure represents a star network and the node in the middle represents the tutor responsible for distributing information and news about the course to the students.

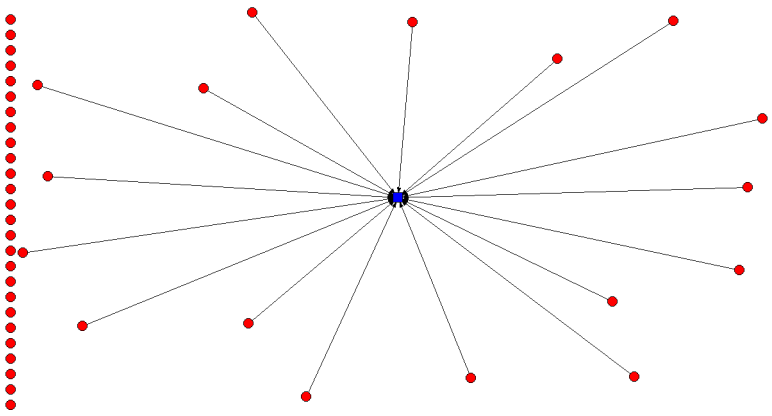


Figure 6.9: Forum's network of course C-2 for the first forum's network

Another example, figure 6.10 is the network of course C-4 for the first forum. The nodes are distributed as small scattered components without a dense component. Moreover, 6.11 shows the sociogram of course C-9 for the first forum, in a circular shape.

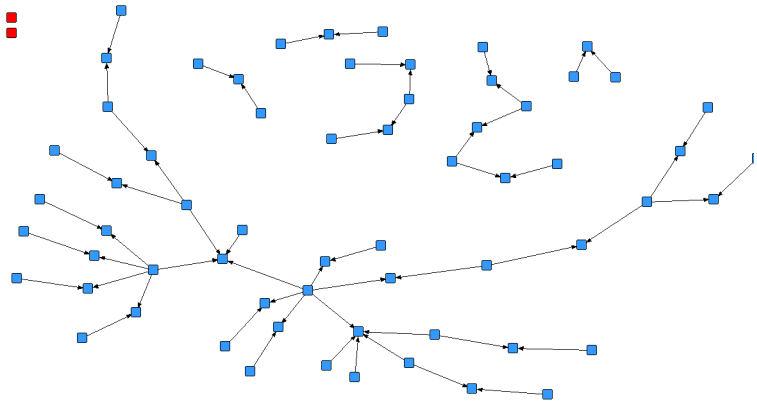


Figure 6.10: Forum's network of course C-4 for the first forum's network

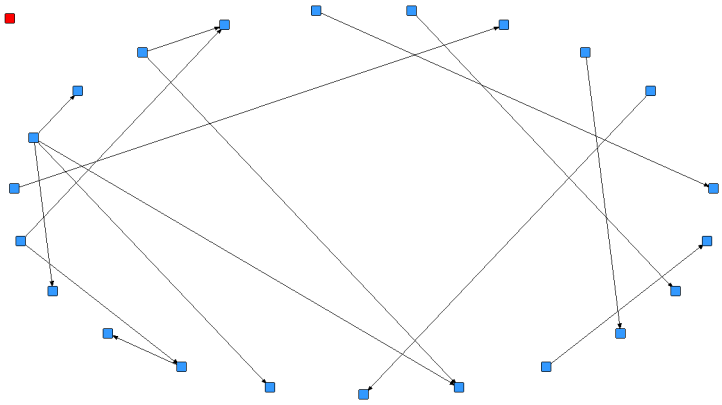


Figure 6.11: Forum's network of course C-9 for the first forum's network

Also, 6.12 is the network of course C-3 for the first forum network. It represents separate dense components and isolated nodes on the left. 6.13 is the network of course C-2.

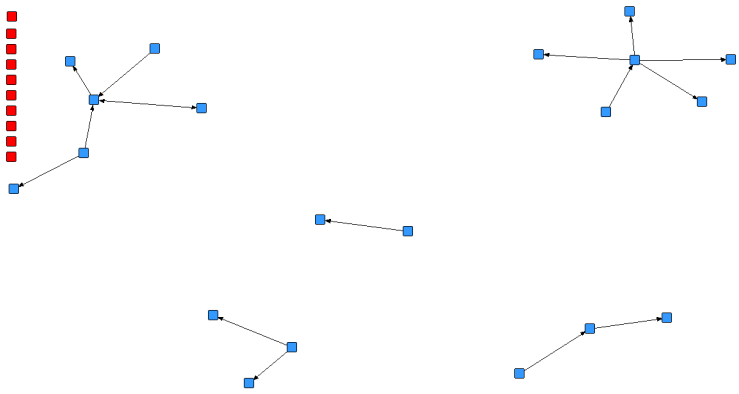


Figure 6.12: Forum's network of course C-3 for the first forum's network

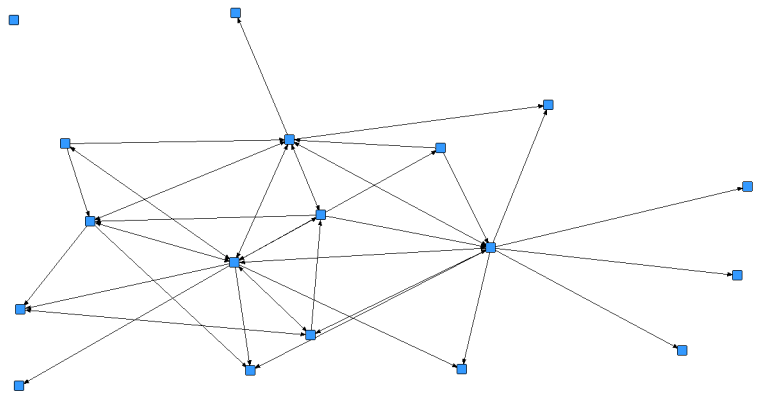


Figure 6.13: Forum's network of course C-2 for the first forum's network

6.3 Wikis' Networks Analysis

The second type of learning network studied are wiki networks. Wikis module within ISIS are still in the early stages of their use. Wiki is a set of webpages that are collaboratively created, edited and shared. Basically, a wiki is a webpage that anyone, depending on its context, can create and modify within the Internet browser. These users need to have minimal HTML basics knowledge. A wiki can be first created with a front page and the authors accordingly add pages to them. The modification and browsing among these pages are done through links since each page is represented as links. Wikis are very useful and important tool as a collaborative environment and encourage individuals to work collaboratively and cooperatively together ((ELI), 2005a).

More specifically, wiki covers two main learning theories constructivism and the collaborative model. Both models differ from the traditional learning approaches, as they both encourage students' participation and empower them to contribute in their learning environment and work jointly with their peers (Kop, 2011).

The following subsections show the definition of wiki networks, as well as the results of studying and applying SNA measures to them. These measures study the structural properties of the network. The different measures applied to wiki networks are those that study the structural position of actors in the network and how it affects the students' performances.

Wiki's Network

The wiki networks are extracted from ISIS's relational database. Wiki modules of 11 courses were analyzed and studied. These specific courses are the ones with the largest number of pages created in their wiki. The wiki's network is defined based on the relation of activities taking place upon wiki pages like creating, editing, or deleting wiki pages for the wikis generated within the course.

For analysis purposes, the log activities of the wikis were saved from ISIS for each course studied. The data was converted into an adjacency matrix in order to represent wiki's pages activities. The definition of the network is defined as an aggregation of wiki's pages activities (i.e. create, modify, or delete) that are performed by either students, teachers, or tutors such as who perform either one of these activities after the other participants

6.3.1 Wikis' Networks Information

11 courses were analyzed and used to extract information for creating wikis' network. Table 6.15 lists the courses along with their names and the number of students registered in each.

Table 6.15: Courses used for wiki's module network

Course	Course Name	Students
C-1	Moodle-Selbstlernkurs	71
C-2	Architektur-Projekt	17
C-3	Architekten-bcp	22
C-4	Globales Projektmanagement	14
C-5	Praktikum zu Grundzägen der Thermodynamik II	20
C-6	BerlinerEnergie Abgedreht	15
C-7	Fachgebietsplattform Maschinen-und Energieanlagentechnik	12
C-8	DGS-FA-HS	
C-9	Energieseminar "Orgaplattform"	30
C-10	GPM+ für Studierende und Jungabsolventen	16
C-11	Technische Grundlagen der Informatik 3	139

Table 6.16 shows the number of pages per each wiki.

The analysis of wiki modules showed that there is no relation between the size of the course and the level of interactions and participation that are taking place within them. This study focuses on the frequency of interactions among students and their willingness to participate in online discussions.

Table 6.17 shows the size of the defined networks. Network size means the number of participants and contributors to a specific wiki. The number of registered students is always larger than the network's size, as all students actively take part in the course.

Table 6.16: Number of pages defined for each wiki

Course	No. of pages
C-1	1
C-2	16
C-3	16
C-4	17
C-5	1
C-6	55
C-7	12
C-8	2
C-9	48
C-10	11
C-11	4

Table 6.17: Network size of wiki' network along with the number of pairs formed

Courses	Wiki Network	
	Size	Pairs
C-1	38	1406
C-2	38	1406
C-3	43	1806
C-4	41	1640
C-5	33	1056
C-6	35	1190
C-7	24	552
C-8	15	210
C-9	13	156
C-10	6	30
C-11	20	380

6.3.2 Cohesion

The density measure is calculated for the networks extracted from wiki's modules along with their participants and activities. Table 6.18 shows mean values of the density measures. The mean values range from 0.025 to 0.1667, they are relatively small. Course C-10 is the one with the highest density value, 0.167, whereas courses C-1 and C-6 have the lowest values, 0.026.

Figure 6.14 shows the distribution of standard deviation values for wiki networks. The standard deviation values range from 0.158 to 0.373. From the figure it can be seen that the variation in the values is less than that of forum networks.

Table 6.18: Density values for wikis networks

Courses	Wiki Network
	Density
C-1	0.026
C-2	0.035
C-3	0.037
C-4	0.041
C-5	0.046
C-6	0.026
C-7	0.044
C-8	0.081
C-9	0.090
C-10	0.167
C-11	0.058

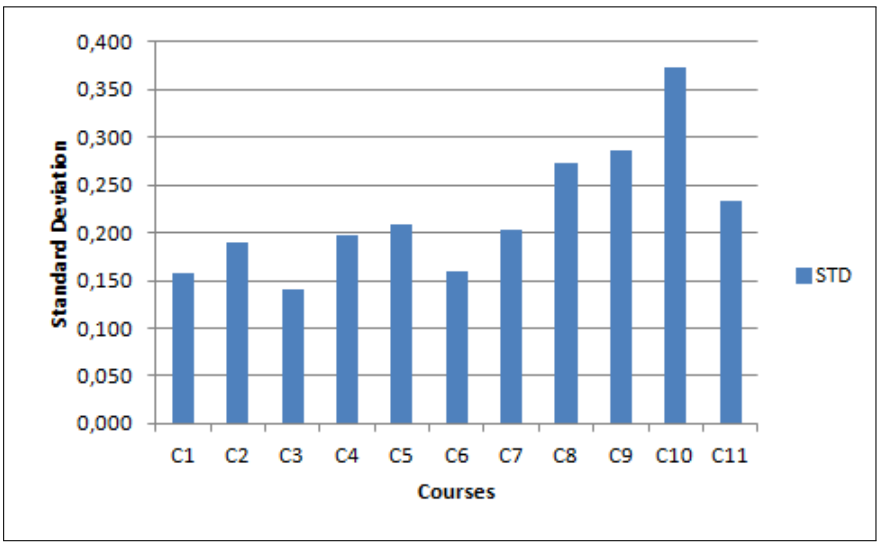


Figure 6.14: Standard deviation values of density measure for wikis' networks

6.3.3 Centrality Measures

Centrality measures are calculated for wiki networks. These measure study the information that can be obtained from the positions of nodes within the network. The next subsection shows the results of the calculation of degree and closeness centrality. The in-betweenness values were all zero for the defined wiki networks.

6.3.3.1 Degree Centrality

Table 6.19 shows mean values of degree centrality for wiki networks. The relations are directed (asymmetric), and therefore two separate degree values were calculated, in- and out-degree. The values ranges from 2.56 to 16.67. The mean values of the forum networks are smaller than the mean values of wiki networks.

Table 6.19: Mean values of degree centrality for wikis' networks

	Wikis' Network	
Courses	Out-degree	In-degree
C-1	2.56%	2.56%
C-2	3.49%	3.49%
C-3	3.71%	3.71%
C-4	4.09%	4.09%
C-5	4.55%	4.55%
C-6	2.61%	2.61%
C-7	4.35%	4.35%
C-8	8.10%	8.10 %
C-9	8.97%	8.97%
C-10	16.67%	16.67%
C-11	5.79%	5.79%

For course C-1, the mean value for both in- and out-degree is 2.56. Course C-10 has the highest mean value of 16.67, the value is the same for both in-degree and out-degree values. The range of mean values of degree centrality is relatively small. This shows that the usage of wikis within these different courses is minimal.

Figure 6.15 shows the standard deviation values of wiki networks. Course C-6 is not much consistent concerning in-degree values. In other words, the number of received messages is higher than the number of sent ones. The network is more centralized around one specific node or individual, who receives all the messages around him. Course C-7 on the other hand has a higher out-degree variance than in-degree variance. The same applies to courses C-8, C-9, and C-10.

Figure 6.15 shows the standard deviation values of degree measure values. A low variance value means that the majority of values are close to the mean value of this network. The values ranges from 1.57 to 20.27. High variance values show how individuals within the network are not consistent in their behavior regarding the defined relation.

Table 6.20 shows the minimum and maximum values for both, the in-degree values and out-degree values. The lowest in-degree (number of messages received) is 0 while the highest values range from 1 to 11 (for course C-2). For out-degree, maximum values range from 1 to 6 (also course C-2).

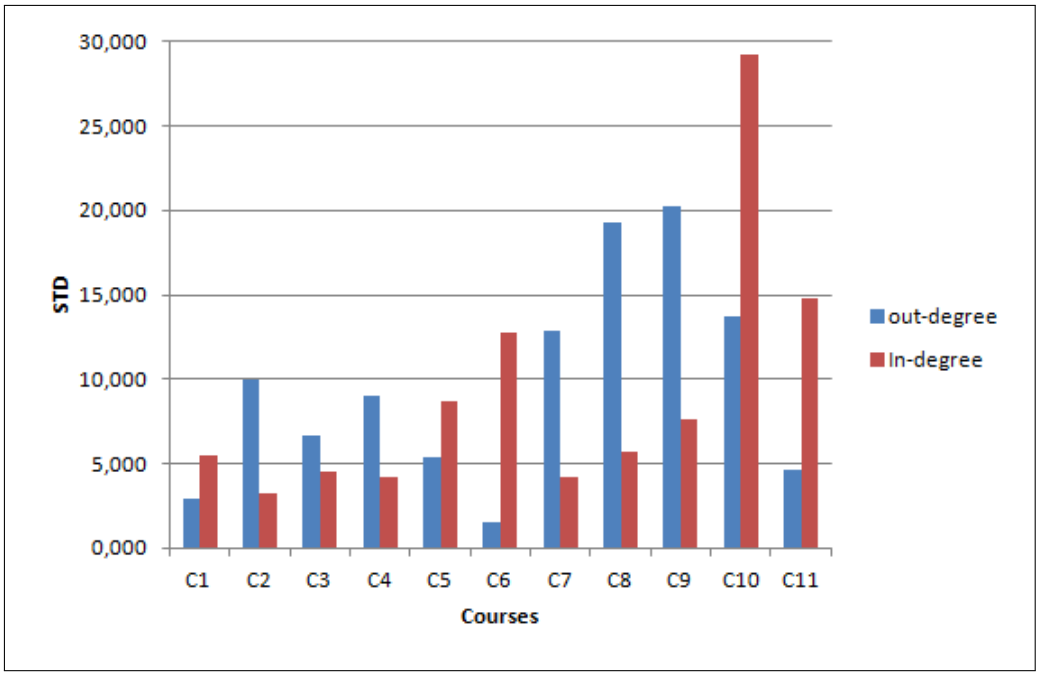


Figure 6.15: Standard deviation values of degree centrality for wikis' network

Table 6.20: Minimum and maximum values of degree centrality for wiki networks

Courses	**Wiki Networks**			
	In-degree		Out-degree	
	Min	Max	Min	Max
C-1	0	1	0	3
C-2	0	11	0	6
C-3	0	4	0	3
C-4	0	6	0	4
C-5	0	1	0	1
C-6	0	1	0	1
C-7	0	3	0	1
C-8	0	1	0	1
C-9	0	1	0	1
C-10	0	3	0	1
C-11	0	1	0	1

Table 6.20 shows the minimum and the maximum number of messages that were exchanged in the network. Obviously, the number of messages exchanged is small. Course C-2 has a maximum number of 11 messages, C-4 had 4 message exchange. Course C-2 shows the largest number of messages exchanged in both, in-degree or out-degree.

6.3.3.2 Closeness Centrality

The closeness measure is considered give an overall meaning of node's centrality since it takes into account the node's or individual's location in the network. The node with highest closeness value means this specific individual is reachable to all other individuals within the network through a minimal distance (Hanneman & Riddle, 2005). Table 6.21 shows the overall mean values of the closeness values for wiki networks.

Table 6.21: Mean values of closeness centrality for wiki networks

	Wiki Network	
Courses	In-closeness	Out-closeness
C-1	2.71%	2.70%
C-2	2.73%	2.77%
C-3	2.42%	2.43%
C-4	2.55%	2.57%
C-5	3.20%	3.18%
C-6	3.11%	2.93%
C-7	4.36%	4.53%
C-8	7.24%	7.71%
C-9	8.44%	9.30%
C-10	22.78%	19.72%
C-11	5.53%	5.30%

The range of closeness values varies from 2.42 to 22.78 for in-closeness and from 2.93 to 19.72 for out-closeness. It is also noticeable that course C-10 has the highest closeness values as well as degree centrality values. Courses C-2, C-3, and C-4 all have small closeness values for both in-closeness and out-closeness. The distribution of in-closeness is less variable than that of out-closeness.

Figure 6.16 shows the distribution of standard deviation values of wiki networks. The variance is very low for courses C-1, C-2, C-3, C-4, and C-5. The largest variance in values is in course C-10.

Another way to study closeness measure values is by presenting the minimum and maximum values of the in-closeness and out-closeness calculated for each node. Table 6.22 shows the minimum and maximum values for closeness centrality. The maximum in-closeness value is 50.0 for course C-10 and the minimum is 2.33 for course C-3. The maximum out-closeness value is 25 for course C-10 and the minimum is 2.33 for course C-3.

Finally, figure 6.17 shows the network of course C-11. The node sizes represent the nodes with higher in-closeness values.

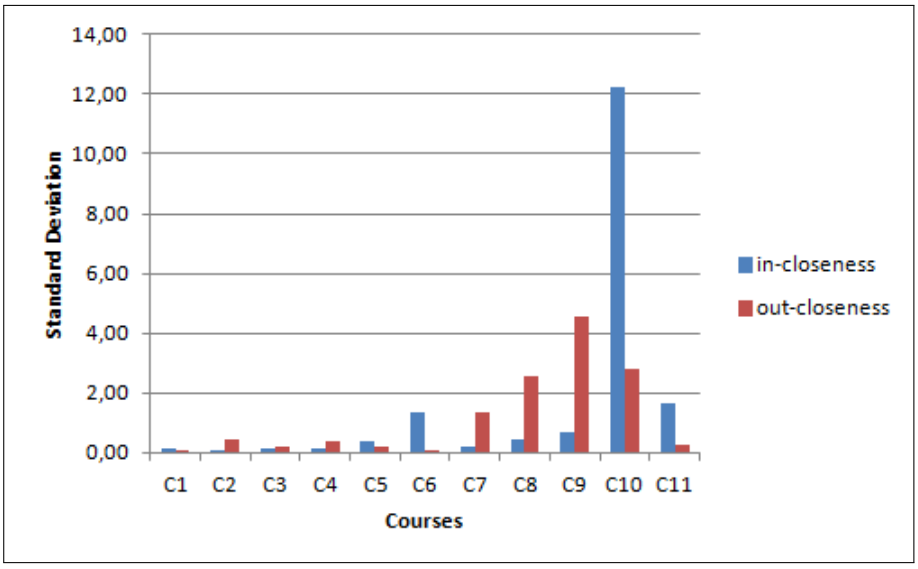


Figure 6.16: Standard deviation values of closeness centrality for wikis' network

Table 6.22: Minimum and maximum values of closeness centrality for wiki networks

	Wiki Network			
Courses	In-closeness		Out-closeness	
	Min	Max	Min	Max
C-1	2.63	3.23	2.63	3.22
C-2	2.63	3.03	2.63	5.0
C-3	2.33	2.86	2.33	3.13
C-4	2.44	2.86	2.44	4.55
C-5	3.03	5.0	3.03	4.17
C-6	2.86	11.11	2.86	3.13
C-7	4.17	5.0	4.17	11.11
C-8	6.67	8.33	6.67	14.29
C-9	7.70	10.0	7.70	25.0
C-10	16.67	50.0	16.67	25.0
C-11	5.0	12.5	5.0	6.25

6.3.4 Centralization

As stated previously, centralization measure computes the overall centrality of the whole network. Its result points out the node or individual that dominates the communication within network. The highly dominate individuals controls all the communication and interactions within the network (Müller-prothmann, 2007). The results of applying this measure for wikis networks range from 0% to 100%. Moreover, they show the difference in the output values in comparison to the star, ring or line pattern.

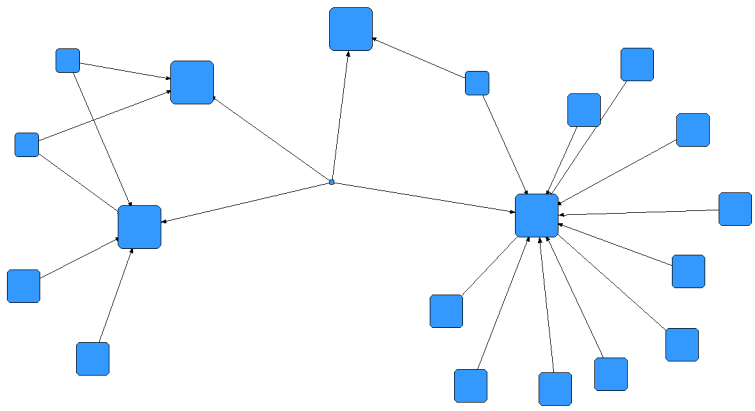


Figure 6.17: Wikis networks for Course C-11, larger node's size represent higher in-closeness values

Table 6.23 shows the network centralization values for both, out- and in-degree centralization. The in-degree values ranges from 10.30 to 76.04 while the out-degree values range from 6.40 to 71.53.

Higher centralization values (for in-degree, and out-degree) mean that there is a substantial amount of concentration (centralization) within the whole network. This means that there is an unequal distribution within the network and that it is centralized around one specific individual, in this case is the tutor.

Table 6.23: Network centralization values for wiki networks for both out-degree and in-degree

	Wiki Network	
Courses	Out-degree	In-degree
C-1	16.80%	16.80%
C-2	46.38%	10.30%
C-3	23.02%	15.70%
C-4	44.50%	11.19%
C-5	24.32%	37.21%
C-6	6.40%	76.04%
C-7	63.52%	13.61%
C-8	52.55%	14.29%
C-9	71.53%	17.36%
C-10	28.0%	76.0%
C-11	16.07%	60.39%

Figure 6.18 shows the distribution of centralization values (both in-degree and out-degree).

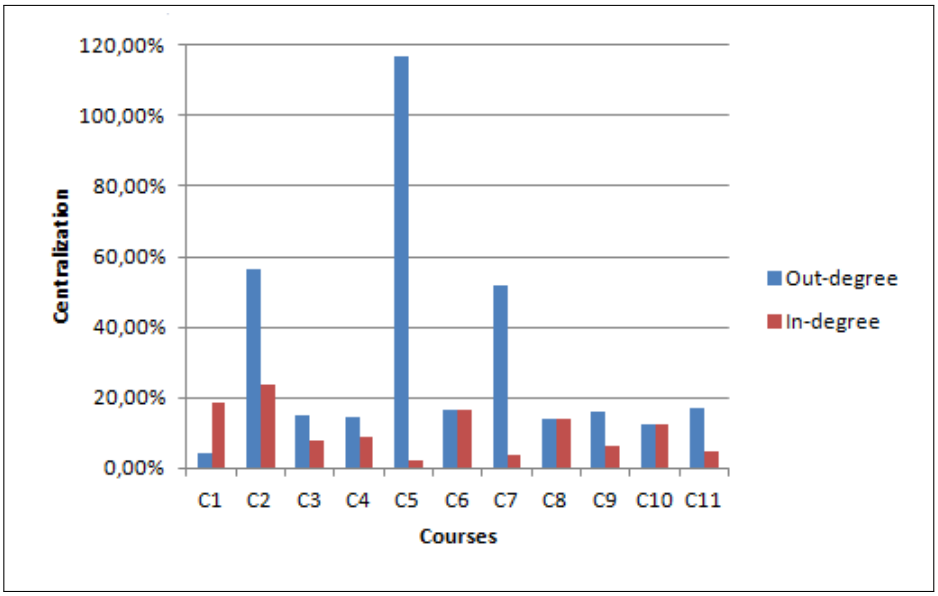


Figure 6.18: Distribution of centralization values for wikis network

6.3.5 Clustering and Subgroups

As explained earlier, studying subgroups, communities, and cliques is considered as a way of studying the individuals tendency to gather and work collaboratively in smaller sets or groups. Throughout this subsection, the measures of clustering coefficients, cliques, and N-cliques are applied to wiki networks.

The result of calculating the clustering coefficient is zeros in all our cases. The greater the clustering coefficient values, the higher the tendencies of finding subgroups (or communities) within the whole network.

6.3.5.1 Cliques and N-Cliques

Clique analysis is the most used measures to identify subgroups within a network. The speed of information distribution is faster among the members of the connected groups (Müller-prothmann, 2007). This measure was applied to the wiki networks, but unfortunately no cliques were found as the clustering coefficient was zero.

Thus, the N-clique method was applied to the wiki networks, with N values of 2, 3, and 5. Table 6.12 shows how many subgroups were found within wikis. Course C-3 has the maximum number of cliques found for all values of N. It has 34 cliques with N equals 2, 28 cliques with N equals 3, and 5 cliques with N equals 5. The minimum number of N-cliques members is 1. The smallest groups were found in course C-10.

Table 6.24: Number of N-cliques formed from wikis network with different N values

Courses	**Wiki's Network**		
	2	**3**	**5**
C-1	8	7	1
C-2	18	10	2
C-3	34	28	5
C-4	31	27	4
C-5	22	13	1
C-6	2	2	2
C-7	6	5	2
C-8	4	2	1
C-9	4	2	1
C-10	2	1	1
C-11	5	2	1

Figure 6.19 represents a sociogram for the N-clique set with an N value of 3, for the wiki network of course C-7. It represents 5 cliques by the blue square, with the members of the cliques (either students, teachers, or tutors) in the red circles.

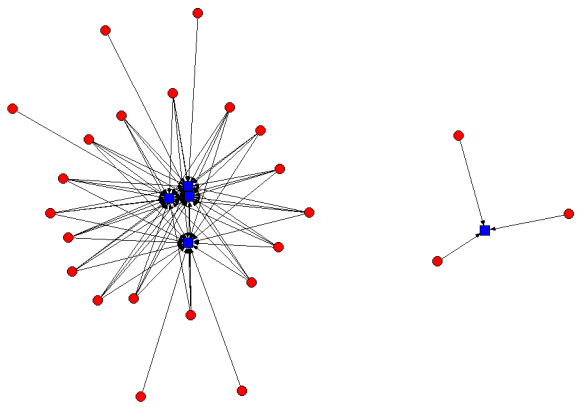


Figure 6.19: N-clique (N is 3) for course C-7 for wiki's network

6.3.5.2 Components

The components measure is also calculated for wikis network. Table 6.25 shows the number of strong and weak components found in the wiki networks. It is noticeable that most of the networks have larger numbers of strong components than weak ones. Course C-3 has the largest number of strong components whereas course C-10 has the lowest number. Generally, the wiki networks show one big component and several small separate components.

Table 6.25: Number of weak and strong components of wikis' network

Courses	**Wiki Network** Weak Comp.	Strong Comp.
C-1	2	38
C-2	1	38
C-3	5	43
C-4	1	41
C-5	1	33
C-6	4	35
C-7	2	24
C-8	1	15
C-9	1	13
C-10	1	6
C-11	1	20

6.3.5.3 Ego-Networks

The ego-network measure is also calculated for the wiki networks. Table 6.26 shows the results of the ego networks of two courses, C-1 and C-6. The egos chosen are the ones with the largest number of neighbors. In course C-1 the ego with the largest number is the teacher and the second largest is the tutor. The egos with the smallest number of neighbors are found to be the students.

In course C-6 the ego with the largest number is that of the tutor. The other egos shown as example are for different students within this course.

Table 6.26: Ego-network measures for courses C-1 and C-6

Ego	Size	Pairs	Density	Broker
Teacher	7	42	0	1
Tutor	5	20	0	1
Student	1	1	0	0

(a) Course C-1 for wiki's network

Ego	Size	Pairs	Density	Brokers
Tutor	26	560	0	1
Student	3	12	0	1
Student	1	1	0	0

(b) course C-6 for wiki's network

6.3.6 Networks Visualization

The following figures represent the most frequently repeating sociograms obtained from the various interactions within the wikis. Unlike for the forum networks, the

ring network pattern was more prevalent. Figure 6.23 shows the sociogram of course C-1, the highest dense wiki network.

Figure 6.20 shows the sociogram of course C-2, with its three main dense components. Moreover, figure 6.21 is the sociogram of course C-4, and shows one densely joint component.

Additionally, figure 6.22 shows the visualization of course C-6. It shows three main separate components, and only one dense components.

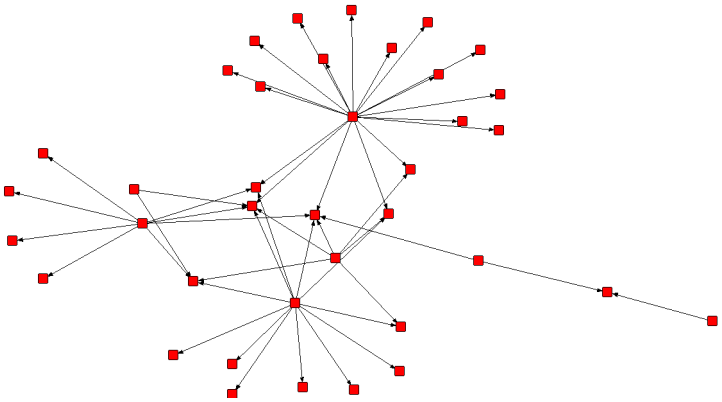


Figure 6.20: Wikis networks for Course C-2

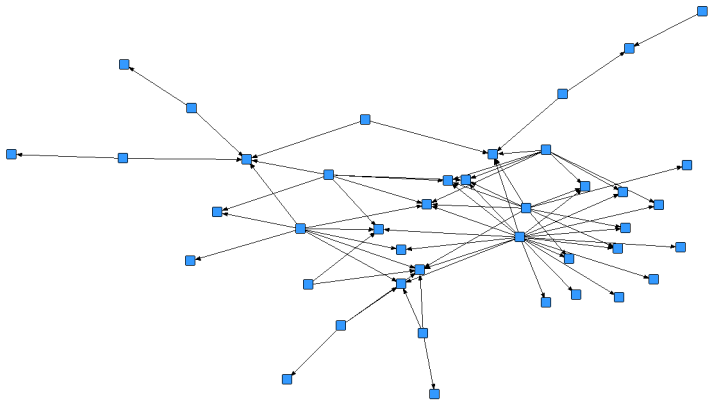


Figure 6.21: Wikis networks for Course C-4

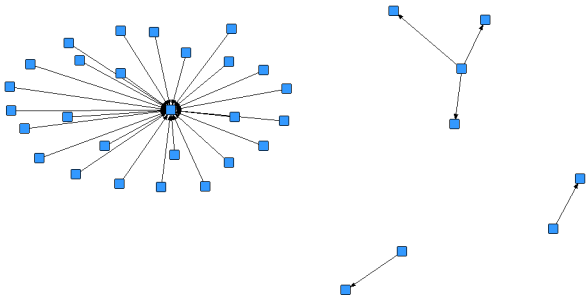


Figure 6.22: Wikis networks for Course C-6

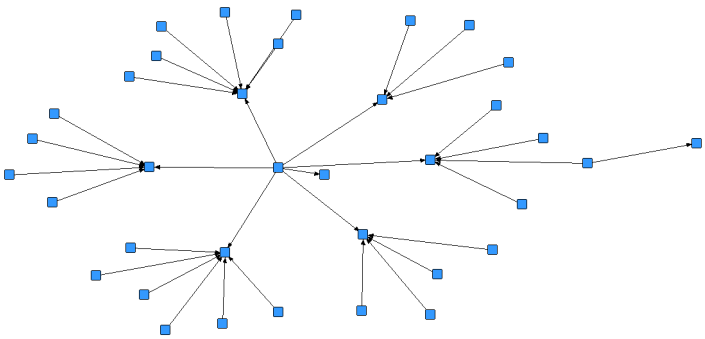


Figure 6.23: Wikis networks for Course C-1

7 Chapter 7 Conclusions and Results

7.1 Introduction

LMSs have largely been used by various learning institutions like universities, schools, or any learning centers. They are mostly used as content management systems where students access to learning material, courses' news, online exams, group applications, and homework submissions. These systems saves all the actions done by either students, teachers, and tutors. These systems store personal information about their users like names, age, year of study, and grades. Large volumes of data are thus generated and saved in relational databases.

However, most of the information produced by LMSs is not being utilized much. They are considered a valuable material for further analysis. It can provide insight into students' behavior and interactions. This provide an overview of the whole dynamics within courses based on the type of networks formulated. For example, it provides information on who is highly active within the class and participates in the different modules of the course. It also allows the analysis of the relation between students' outcomes and level of participation, or of the effect of a teacher's presence on participation.

This thesis studied learning networks extracted and formed from students' interactions within the LMS. LMSs have different interactive modules such as discussion forums, wikis, chat messages, assignments, quizzes, instant messaging, and e-mails. This study focuses on the networks extracted from discussion forums and wikis, as these have collaborative and cooperative aspects. These modules offer a place where students can share and exchange ideas with their classmates.

SNA is used as an analytical tool to study the extracted networks and their characteristics. The analysis is based on two main approaches; mathematical and visual analysis. Moreover, SNA as a research methodology identifies basic network properties, positions of network members, characteristics of relations, cohesive sub-groups, and cutpoints that controls flow of information. These measures were applied to both discussion forums and wiki networks.

After studying and analyzing these networks, many observations can be understood about courses' internal communication patterns. For example, teachers can offer new approaches for those students who are less active and participative in the course and have poor outcomes. To motivate students to be more active in the courses' modules like wikis, discussion forums, messages, or chats can be used. At least, teachers need to assign higher priorities for discussion forums and encourage tutors to be more active in using the different learning modules in the LMS. Teachers can force these interactions to take place by associating grades or additional points as a bonus to the final grades to it. The presence of the teacher within these modules also motivates their students to highly participate.

Additionally, it is necessary to motivate students to participate more in wikis, for example publishing news and discussions about the final year project in them. It is considered a better assessment of the participation of learners within the course. This information allows teachers to change the teaching policy within the class or to initiate learning groups as part of the teaching practice. Teachers should find a way to combine students from different grade levels into one group. It is more likely to foster dynamic communication in the subsequent activities.

Finally, teachers need to assign higher priorities to modules within LMS that can encourage students to use them and always turn back to their classmates for solving problems. Students could get more focused comments and advice about their problems to help them with their problems in the projects they have to present, for example a final year or master projects.

Recently, research areas about SNA seem to have gained much more interest, especially with the great expansion of social networks like Facebook. Learning networks and analytics however are still in the early stages of the research process and offer larger number of topics for further discussion.

7.2 Summary of Findings

This study uses SNA as an analytical tool to study the different interactions of students within learning environments. SNA offers two main approaches; mathematical analysis and visualization, i.e. sociograms. The analysis results can help teachers and/or tutors to interpret the social structure of the networks extracted from activities of learning modules within the activities and thus gain insight on students' participation patterns and styles in the course.

This study also focuses on the structural properties of networks such as density, centralization, centrality, and clustering, and how these are related to the performance of the students. In general, SNA enables researchers to understand how individuals are connected within a network. Moreover, group density and cohesion is measured to examine group's characteristics and formation within a network.

Finding subgroups, cliques, and components is the first step to finding communities within courses.

The objective of this study was to present information that could be useful for the teachers about the students' performance, actions, and interactions within the course. The next subsections 7.2.1 and 7.2.2 provide a summary of the main analytical findings.

7.2.1 Thoughts about Forum's Networks

Discussion forums is one of the modules offered by Moodle, where students can exchange information during the course. In this study, 11 courses were analyzed to extract and later analyze the underlying networks. These courses defined many forums, but not all of them were used, there were not many participants. This results in very small forum networks. The difference in networks' sizes depends on the nature of the forum. Different forums were defined in these courses, for example news forums to distribute information about the course, forums to submit assignments or forums to submit questions for tutors, and teachers regarding the course.

The forums selected for the purpose of analysis are the ones with the most participants. The densities of these networks were very small, giving an indication of the low participation in these forums. Forum networks with higher densities were the ones related to either assignment submission or question asking ones (in which the teachers participate). In the question forums the larger number of posts were created by tutors.

Many aspects of network measures were studied. These different measures help to understand the relation between an actor's position in the network and the main characteristics of the students within the course. For example, degree centrality points out the most and less influential actors within the network. The results were almost similar for all courses actors with higher degrees are the tutors. In other courses the student who submitted the maximum number of assignments were the same as those that got better grades. Ironically, some courses did not show any relation between the importance of the student within the network and his/her grade outcomes.

The number of interactions measured for the various forums depends on the type of the forum and on how frequently students use them. It was difficult to determine the exact role of teachers within these forums. Larger network size does not result in a denser network. This is because most of the forums were just news forums to distribute information within the class or homework submitting forums for students. The courses where the students have the highest degree values were the discussion forums about submitting homeworks.

Students have a very low tendency to interact within the forum networks and to exchange ideas. These networks extracted from modules of LMS are considered as formal learning networks, students prefer to communicate through informal networks like Facebook or social network services. Increasing students' participation level has to be enforced by teachers or tutors and to be linked with their final grades.

The results of centralization measure shows that most of the networks have either the tutors and sometimes the teachers as the main players within the network. Some networks have students as the central node, but those were the students who submitted the largest number of assignments.

As a result of low density values, betweenness values were extremely small. This type of centrality is another indicator of how highly connected students are, and less participative. Low values mean students are highly independent and not willing to group with others to work jointly. This further emphasizes the idea of very weak participation in the forums.

Closeness centrality was also calculated for forum networks. This measure was calculated for the two networks. Closeness values for the second network were slightly higher than those of the first network in almost all cases. The overall closeness values of the whole network could not be calculated since the network has many disconnected components. This is applied for the two defined networks. This is another indication of the weak participation, low active students, and very low connectivity within the network. All of these obtained conclusions from the networks give an overall understanding and an important feedback of the class dynamics.

The second part of the application of the measures was related to the clustering measure and to the subgroups i.e. cliques. These measures gives an overview of how much students tend to collaborate with their peers and work together in groups. This can be measured by applying clustering coefficient measure to the whole networks. This measure was calculated for forum networks; the result was zero for the majority of the forum networks. This is interpreted by the less tendency of students to work in groups. Thus, it was difficult to find smaller groups i.e. cliques within these networks. The need was to find smaller group sizes by measuring the N-cliques. N-clique is a less rigid definition for detecting groups within networks. Smaller groups sizes where detected only through smaller N sizes from N-cliques.

The previous findings can be explained by the very weak participation within these forums. It was also detected that the members of the different groups are similar. Students usually gathered around tutors or teachers. Only cliques with larger sizes belong to forums where tutors were assigned to answer students' questions. On the other hand, cliques which were found in homework forums belong of students with more or less similar academic behavior.

Finding components within forum networks is considered as another approach for understanding the tendency among students to work collaboratively together. There are two types of components, they are strong and weak components. The difference between defining them is based on the direction of relations among their members. Basically, the number of components, either strong or weak, was larger for the first network than for the second one. There were generally more strong than weak components. Moreover, it is found that forum networks consisted mainly of several disconnected, small components. There were no big components found, only many isolated members who did not participate. The isolated students were the ones who did not submit any homework. All of these findings emphasize the idea that students are less active in the class and not willing to engage in communications through modules of LMS.

The second part of analyzing learning networks extracted from the discussion forums is done by visualizing these networks through sociograms. A repeated pattern found in the forum network was the star network pattern. This pattern represent the network member who is the most centralized node in the network. This centralized individual controls all the communications within the network as well as controlling all the messages sent and received between the other members. The network connectivity in this pattern depend on the central node.

Finally, summing up all these previous results shows that learning networks were almost not created within the courses. The size of these networks are very small due to the very low participative levels among students. Students are not willing to engage in communication through learning modules, and they prefer more to either meet face-to-face or through informal social networks. Thus a teacher's understanding of the interactions among the learning networks is different of the real interactions that are taking place among students.

7.2.2 Thoughts about Wikis' Network

Wiki networks are the second type of networks extracted and analyzed from ISIS. A wiki is a shared interactive documents where different actors, i.e. students, teachers, and tutors can collaboratively interact. 11 courses were analyzed. Generally, students' participation is very weak within this kind of module. Wiki modules were not recognized as part of the learning strategy within courses. It had to be enforced by teachers, at least at the beginning, to motivate students to use it.

The size of wiki networks were extremely small; it ranged from 6 to 43. These are considered small networks, even though they are slightly larger than forum networks. This is due to the different type of interactions in wikis not only sending and receiving messages but also creating, editing, and modifying wiki pages. SNA measures were also applied to wiki networks. The density measure calculated for these networks was noticeably small.

Centrality measures were also applied to wiki networks. It was noticeable that the nodes with the higher degree values were the ones for the tutors. The highest mean value of the degree measure showed the course with the smallest size. This is explained by the fact that the smaller the course's size, the larger the number of students motivated to participate in the course's modules. The personal relations among small courses is very high and that motivates students to publish and generate wikis for different themes.

Betweenness centrality was zero for all networks, meaning that students are less dependent on their peers to start any communication. In other words, this is a result of less collaboration among students. The speed of information sharing among students are extremely weak. Students receive information more or less directly from the main sender directly to them without depending on "in-between" nodes on the way. This explains the relatively small centralization values of wiki networks in comparison to the forum networks. There is no single node controlling the network communications.

Another centrality measure calculated was closeness centrality. The overall closeness values for the networks were small. However, the in-closeness values for course C-10 was the highest. This is explained by the idea that distance between one specific node and all the others within the network is relatively small. This means that this specific individual can get the new information relatively faster than their neighbors within the network.

Regarding studying groups within learning networks extracted from wikis. Clustering coefficient and cliques measures are calculated for these learning networks. The clustering coefficient values for all the networks were zero, due to the low density values of the network and weak participation of students in wikis.

Consequently, the number of subgroups found in wiki networks is larger than that of forum networks. The sizes of the computed subgroups were larger than those of the forum networks due to the higher densities comparing to forum networks. The higher densities of wikis networks are higher due to the collaborative nature of wiki with is formed from different types of interactions. The tutors' presence was a main factor in the entire subgroup's definition. This also emphasizes the previous findings that students are willing to participate more in these modules once the teacher or tutor are participating as well.

Strong and weak components were also computed in order to understand the ways in which students tend to collaborate and work together. The number of strong components was much larger than that of the weak components. This is also noticeable in the sociograms of wiki networks which represent the visualization of wiki's interactions. They mostly presented as one large, dense component along with one small, sparse one. Another common pattern of wiki networks was the ring network pattern. Basically, this pattern assign an equal importance or priority for all the nodes within the network.

After comparing the two module networks, discussion forums and wikis, it became apparent that students are more familiar with, and more willing to participate in discussion forums than wikis. Discussion forums are mostly used for submitting homeworks or distributing the course news. Other types of forums which are highly used among students where the ones that are created for asking questions or registering for exams.

However, wikis are still in the early stages of development and used very little in courses. This module is relatively new to Moodle and still not used much. Wikis are managed by the tutors of the course, used as a tutorial and place to summarize the main concepts of learning materials. It is used as a place to share, regulate, and dismiss information about learning projects.

7.3 Challenges and Limitations

The main obstacle faced throughout this study was data privacy. This study dealt with students' information and accessed the database about their grades, assignments, and their information. A statement declaring the pure scientific use of this information was signed.

Choosing which course to analyze was done by selecting the ones in which learning modules i.e. discussion forums, wikis, and assignments were most used. The courses included in this study were chosen among many analyzed ones, in order to present the ones containing useful information about students' participation style. It was observed that the courses which use larger numbers of learning modules were the ones with a smaller number of registered students. On the other hand, general courses of the first year, with more than 500 students, have a minimum number of student interactions in ISIS.

There is no direct relation between a course's size, i.e. number of registered students, and the size of the network, i.e. the number of nodes in the network. Not all the registered students in the course are willing to participate in the learning modules online. Moreover, larger courses tend to count few interactions between their students. In these courses, students deal with ISIS only as a content management system.

On the contrary, smaller courses tend to have more important numbers of interactions. Students in these courses tend to participate more in the learning environments since they know each other personally. The frequency of the interactions also varied between courses.

It can be concluded that students are still using the LMS as a repository for learning material. Students are not motivated to refer to the LMS as to a place to find solution to their learning problems, instead they tend to meet face-to-face with their immediate close friends.

7.4 Future Direction

Recently, applying SNA to learning environments and learning problems has attracted a large amount of interest. Different research approaches have been introduced such as learning analytics, networked learning communities, social learning analytics, and many more. Social media applications have an enormous effect on how learning and teaching are taking place nowadays, as can be seen in online courses, MOOC, blogs, and most famously social in network sites, i.e. social media. These applications provide different informal possibilities where students can find solutions collaboratively with their peers.

There are two main future directions. First, analyzing the content of the discussion forums within ISIS discussion forums. This will allow the study of other dimensions of forums' network such as: what is the focus of the network? what are the themes most likely to be discussed?, and what are the factors that affect the formation of learning subgroups in forum networks? Another approach will be defining a new type of relations for forum networks, in other words two-mode network. Defining two-mode network will allow studying the relations between the topic discussed and the students' personal profiles at the same time. Studying two-mode networks will also allow understanding the learning communities characteristics and approaches formulated accordingly.

Additionally, much useful information about students' interaction patterns and their change over time can be obtained through the analysis and visualization of the wiki interactions. Also, analyzing social relationships defined between students and tutors specifically and how they changed over time, i.e. the beginning, middle, or end of the course duration.

It is important to integrate wikis more into the learning activities which rely largely on the collaboration between students. One of the ideas is to allow students to review and comment on their peers work. This will highly encourage students to work together. Another possibility for highly integrating wikis within the class is presenting personal learning plans and project proposals. This helps students exchange valuable ideas and finding solutions for any problems encountered throughout their studies.

Another important aspect to be further analyzed is the comparison of the participation between two guided networks. These two networks are formal learning networks (i.e. extracted from learning modules of LMS) and informal learning network (i.e. extracted from MOOC courses, student's e-portfolios, and PLE). Analyzing learning subgroups that are formed from informal sources for learning networks and what are the factors that affect the formation of these networks and the detection of communities and subgroups within them. Moreover, analyzing the differences between forming the informal and formal networks.

Bibliography

Abello, J., Pardalos, P.M., & Resende, M.G.C. 1998. On very large maximum clique problems. *Pages 175 – 183 of:* Battiti, R., & Bertossi, A. (eds), *Proceedings of algorithms and experiments (ALEX98).*

Aiello, William, Chung, Fan R. K., & Lu, Linyuan. 2000. A random graph model for massive graphs. *Pages 171 – 180 of:* Yao, F. Frances, & Luks, Eugene M. (eds), *STOC*. ACM.

Albert, Réka, & Barabási, Albert-László. 2002. Statistical mechanics of complex networks. *Reviews of modern physics*, **74**(January), 78.

Alias, Nor Aziah, & Zainuddin, Ahmad Marzuki. 2005. Innovation for better teaching and learning: Adopting the learning management system. *Malaysian online journal of instructional technology*, **2**(2), 27 – 40.

Amaral, L. A. N., Scala, A., Barthélémy, M., & Stanley, H. E. 2000. Classes of small–world networks. *Proceedings of the national academy of sciences*, **97**(21), 11149 – 11152.

Amaratunga, D., Baldry, D., Sarshar, M., & Newton, R. 2002. Quantitative and qualitative research in the built environment: Application of "mixed" research approach. *Work study*, **51**(1), 17 – 31.

Anderson, Paul. 2007. What is Web 2.0? ideas, technologies and implications for education. *Technology*, **Feb**(1), 1 – 64.

Anderson, T., & Elloumi, F. 2004. *Toward a theory of online learning, theory and practice of online learning*. Athabasca University.

Aspin, David N, & Chapman, Judith D. 2000. Lifelong learning: concepts and conceptions. *International journal of lifelong education*, **19**(1), 2 – 19.

Attwell, Graham. 2007a. Personal learning environments – the future of elearning ? *Lifelong learning*, **2**(January), 1 – 8.

Attwell, Graham. 2007b. Personal learning environments for creating, consuming, remixing and sharing. *Mind*, 36 – 41.

Aviv, Reuven, Erlich, Zippy, Ravid, Gilad, & Geva, Aviva. 2003. Network analysis of knowledge construction in asynchronous learning networks. *Analysis*, **7**(3), 1 – 23.

Baker, Ryan, & Yacef, Kalina. 2009. The state of educational data mining in 2009: A review and future visions. *Review literature and arts of the americas*, **1**(1), 3 – 17.

Barabási, Albert-László, & Bonabeau, Eric. 2003. Scale–free networks. *Scientific american*, **288**(5), 50–59.

Barneveld, Angela Van, Arnold, Kimberly E., & Campbell, John P. 2012. Analytics in higher education: Establishing a common language(ELI paper 1). *EDUCAUSE learning initiative*, **1**(January), 1 – 11.

Bastian, Mathieu, Heymann, Sebastien, & Jacomy, Mathieu. 2009. Gephi: An open source software for exploring and manipulating networks. *In: International AAAI conference on Weblogs and social media.*

Batagelj, Vladimir, & Mrvar, Andrej. 1998. Pajek – program for large network analysis. *Connections*, **21**, 47 – 57.

Bates, Tony. 2010. Understanding Web 2.0 and its implications for e–learning. *In Web 2.0–based e-learning: Applying social informatics for tertiary teaching*, 21 – 42.

Berlanga, AJ, Rusman, E., Eshuis, JH, Hermans, H., & Sloep, PB. 2009. Learning networks for lifelong learning: An exploratory survey on distance learners' preferences. *In: Proceedings of the 7th international conference on networked learning.*

Bertini, Enrico. 2005. *Social network visualization: A brief survey.* [Accessed: February 2014].

Bessenyei, I. 2008. Learning and teaching in the information society. elearning 2.0 and connectivism. *Revista de informatica sociala*, **2008**(9), 1 – 11.

Bienkowski, Marie, Feng, Mingyu, & Means, Barbara. 2012. Enhancing teaching and learning through educational data mining and learning analytics: An issue brief. *In: Office of educational technology. U.S. department of education.*

Black, Erik W., Dawson, Kara, & Priem, Jason. 2008. Data for free: Using LMS activity logs to measure community in online courses. *Pages 65 – 70 of: The internet and higher education*, vol. 11.

Borgatti, S P, Everett, M G, & Freeman, L C. 2002. UCINET for windows: Software for social network analysis. *Harvard analytic technologies*, **2006**.

Borgatti, S P, Everett, M G, & Freeman, L C. 2010. *Analytictech – UCINET.* `http://www.analytictech.com/ucinet/description.htm`. [Accessed: February 2014].

Borgatti, Stephen. 1995. Centrality and AIDS. *Connections*, **18**(1), 112 – 114.

Borgatti, Stephen P, & Foster, Pacey C. 2003. The network paradigm in organizational research: A review and typology. *Journal of management*, **29**(6), 991 – 1013.

Borgatti, Stephen P, Mehra, Ajay, Brass, Daniel J, & Labianca, Giuseppe. 2009. Network analysis in the social sciences. *Science*, **323**(5916), 892 – 895.

Brown, John Seely, & Adler, Richard P. 2008. Minds on fire: open education, the long tail, and learning 2.0. *EDUCAUSE review*, **43**(1), 16 – 32.

Budka, P., & Mader, E. 2006. E–learning in the social sciences: experiences with creating communities, networks and strategies. *In: Learning communities conference, klagenfurt, austria: Klagenfurt university, 9-11 november.*

Burt, Ronald S. 2004. Structural holes and good ideas. *American journal of sociology*, **110**(2), 349 – 399.

Campbell, Aaron, Ammann, Rudolf, & Dieu, Barbara. 2005. Elgg – a personal learning landscape. *Teaching english as a second or foreign language TESL–EJ*, **9**(2).

Campbell, J P, DeBlois, P B, & Oblinger, D G. 2007. Academic analytics. *EDUCAUSE review*, **42**(2), 40 – 57.

Carugati, F., Selleri, P, Matteucci, M., Tomasetto, C., Mazzoni, E., & Gaffuri, P. 2008. Monitoring and analyzing collaboration in e–learning environment: two case studies applied to vocational training. *In: Conference on knowledge construction in e–learning context: CSCL, ODL, ICT and SNA in education.*

Castro, F., Vellido, Alfredo, Nebot, A., & Mugica, F. 2007. Applying Data Mining Techniques to e-Learning Problems. *Chap. 8, pages 183 – 221 of:* Kacprzyk, Janusz, Jain, Lakhmi, Tedman, Raymond, & Tedman, Debra (eds), *Evolution of teaching and learning paradigms in intelligent environment.* Studies in Computational Intelligence, vol. 62. Berlin, Heidelberg: Springer Berlin Heidelberg.

Catanese, Salvatore, De Meo, Pasquale, Ferrara, Emilio, Fiumara, Giacomo, & Provetti, Alessandro. 2012. *Extraction and Analysis of Facebook Friendship Relations.*

Chatti, M.A., Dyckhoff, A.L., Thüs, H., & Schroeder, U. 2012. A reference model for learning analytics. *International journal of technology enhanced learning IJTEL*, **4**(5 – 6), 318 – 331.

Chatti, Mohamed Amine. 2010. *Personalization in technology enhanced learning.* Ph.D. thesis, Mathematic and Computer Science RWTH Aachen Germany.

Chatti, Mohamed Amine, Jarke, Matthias, & Frosch-Wilke, Dirk. 2007. The future of e–learning: a shift to knowledge networking and social software. *International journal of knowledge and learning IJKL*, **3**(4/5), 404 – 420.

Cho, H, Gay, G, Davidson, B, & Ingraffea, A. 2007. Social networks, communication styles, and learning performance in a CSCL community. *Computers & education*, **49**(2), 309 – 329.

Cho, Hichang, Stefanone, Michael, & Gay, Geri. 2002. Social information sharing in a CSCL community. *Proceedings of the conference on computer support for collaborative learning foundations for a CSCL community CSCL 02*, **28**(7), 43.

Christakis, Nicholas A, & Fowler, James H. 2007. The spread of obesity in a large social network over 32 years. *The new england journal of medicine*, **357**(4), 370 – 379.

Cohen, Don, & Prusak, Laurence. 2001. *In good company: How social capital makes organizations work*. Vol. 15. Harvard Business School Press.

Cole, C., & Foster, H. 2007. *Using moodle*. `http://issuu.com/iusher/docs/usingmoodle2`. [Accessed: February 2014].

Conole, G., de Laat, M., Dillon, T., & Darby, J. 2008. 'disruptive technologies', 'pedagogical innovation': What's new? findings from an in–depth study of students' use and perception of technology. *Computers & education*, **50**(2), 511 – 524.

Coulon, Fabrice. 2005. The use of social network analysis in innovation research: A literature review. *In: DRUID academy winter 2005*.

Craik, F, & Lockhart, R. 1972. Levels of processing: A framework for memory research. *Journal of verbal learning and verbal behavior*, **11**(6), 671–684.

Creswell, John W. 2002. *Research design: Qualitative, quantitative, and mixed methods approaches*. SAGE Publications.

Cross, Jay. 2006. *Informal learning: rediscovering the natural pathways that inspire innovation and performance*. 1 edn. Pfeiffer Wiley.

Dagger, Declan, O'Connor, Alexander, Lawless, Séamus, Walsh, Eddie, & Wade, Vincent P. 2007. Service–oriented e–learning platforms: From monolithic systems to flexible services. *IEEE internet computing*, **11**(3), 28 – 35.

Dalcher, Darren. 2004. Stories and histories: case study research (and beyond) in information systems failures. *Chap. 32, pages 305 – 322 of:* Klosgen, W., & Zytkow, J. (eds), *The handbook of information systems research*. Hershey, PA: Idea Group Publishing.

Daniel, Ben K, McCalla, Gordon I, & Schwier, Richard A. 2008. Social network analysis techniques: implications for information and knowledge sharing in virtual learning communities. *International journal of advanced media and communication*, **2**(1), 20.

Daradoumis, Thanasis, Martínez-Monés, Alejandra, & Xhafa, Fatos. 2004. An integrated approach for analysing and assessing the performance of virtual learning groups. *Pages 289 – 304 of: Groupware: Design, implementation, and useCRIWG*.

Dawson, Shane. 2006. Online forum discussion interactions as an indicator of student community. *Australasian journal educational technology AJET 22*, **22**, 495 – 510.

Dawson, Shane. 2008. A study of the relationship between student social networks and sense of community. *Educational technology & society*, **11**(3), 224 – 238.

Dawson, Shane, McWilliam, Erica, & Tan, Jen Pei-Ling. 2008. Teaching smarter: How mining ICT data can inform and improve learning and teaching practice. *Pages 221 – 230 of: In hello! where are you in the landscape of educational technology? proceedings, ASCILITE: Melbourne, australia.*

De Laat, M, Lally, V, Lipponen, L, & Simons, R. 2007. Investigating patterns of interaction in networked learning and computer–supported collaborative learning: A role for social network analysis. *International journal computer–supported collaborative learning*, **2**(1), 87 – 103.

de Nooy, Wouter. 2009. Social network analysis, graph theoretical approaches to. *Pages 8231 – 8245 of: Encyclopedia of complexity and systems science.* Springer.

de Nooy, Wouter, Mrvar, Andrej, & Batagelj, Vladimir. 2005. *Exploratory social network analysis with pajek (structural analysis in the social sciences).* Cambridge University Press.

Dekker, Anthony H. 2002. A category – theoretic approach to social network analysis. *Electronic notes in theoretical computer science*, **61**(1 – 5), 21 – 33.

Dougiamas, M. 2007. *Moodle – philosophy.* `http://docs.moodle.org/24/en/Philosophy`. [Accessed: February 2014].

Downes, S. 2010. *Collaboration, analytics, and the LMS: A conversation with stephen downes, campus technology.* `http://campustechnology.com/newsletters/ctfocus/2010/10/collaboration_analytics_and-the-lms_a-conversation-with-stephen-downes.aspx`. [Accessed: February 2014].

Downes, Stephan. 2005. *E–Learning 2.0.* `http://elearnmag.acm.org/featured.cfm?aid=1104968`. [Accessed: February 2014].

Downes, Stephan. 2006. *Learning networks and connective knowledge.* `http://it.coe.uga.edu/itforum/paper92/paper92.html`. [Accessed: February 2014].

Dron, J. 2006. Pedagogy 2.0. *Pages 27 – 32 of: In proceedings of the 12th NETTIES international conference – the future of E: Advance educational technologies for a future e–Europe Timisoara, Romania.*

Dron, Jon, & Anderson, Terry. 2009. Lost in social space: Information retrieval issues in Web 1.5. *Journal of digital information*, **10**(2), 1 – 12.

Dunaway, Michelle Kathleen. 2011. Connectivism: Learning theory and pedagogical practice for networked information landscapes. *Reference services review*, **39**(4), 675 – 685.

Dunne, Cody. 2011. *NodeXL: Network overview, discovery and exploration for excel.* `http://nodexl.codeplex.com/`. [Accessed: February 2014].

Ebner, Martin. 2007. E–Learning 2.0 = e–Learning 1.0 + Web 2.0? *Pages 1235 – 1239 of: Availability, reliability and security, 2007. ARES 2007. the second international conference on.*

(ELI), EDUCAUSE Learning Initiative. 2005a. 7 things you should know about wikis. [Accessed: February 2014].

(ELI), EDUCAUSE Learning Initiatives. 2005b. 7 things you should know about blogs. *EDUCAUSE*, **April**.

Erlin, Yusof, N., & Rahman, A.A. 2009. Students' interactions in online asynchronous discussion forum: A social network analysis. *Pages 25 – 29 of: International conference on education technology and computer (2009) ICETC '09.*

Ferguson, R. 2012. *The state of learning analytics in 2012: a review and future challenges. technical report KMI–12–01.* Tech. rept. Knowledge media institute, the open university, UK.

Freeman, Linton C. 2004. *The development of social network analysis: A study in the sociology of science.* Empirical Press.

Garton, Laura, Haythornthwaite, Caroline, & Wellman, Barry. 1997. Studying online social networks. *Journal of computer–mediated communication*, **3**(1), 0 – 0.

Gerhard, J., Mayr, P., & Seufert, S. 2002. Classroom component of an online learning community: Case study of an MBA program at the university of st. gallen. *In m. khosrow-pour (ed.), web based instructional learning. Hershey: IRM press.*

Giani, Umberto. 2010. E–learning, social networks analysis and the diffusion of innovation. *Journal of eLearning and knowledge society english version*, **6**(1), 51 – 58.

Granovetter, Mark S. 1973. The strength of weak ties. *American journal of sociology*, **78**(6), 1360 – 1380.

Gretzel, Ulrike. 2001. *Social network analysis: Introduction and resources.* `http://lrs.ed.uiuc.edu/tse-portal/analysis/social-network-analysis/`. [Accessed: February 2014].

Grosseck, Gabriela. 2009. To use or not to use Web 2.0 in higher education? *Procedia – social and behavioral sciences*, **1**(1), 478 – 482.

Gulati, Ranjay. 1995. Social structure and alliance formation patterns: A longitudinal analysis. *Administrative science quarterly*, **40**(4), 619 – 652.

Gulati, S. 2003. Informal learning: Building an argument for inclusive online learning. *Athens institute of education and research ATINER*, **6**.

Hamulic, Indira, & Bijedic, Nina. 2009. Social network analysis in virtual learning community at faculty of information technologies. *Procedia social and behavioral sciences*, **1**(1), 2269 – 2273.

Han, Jiawei, Kamber, Micheline, & Pei, Jian. 2006. *Data mining: Concepts and techniques*. 2rd edn. Morgan Kaufmann Publishers Inc.

Hanna, Margo. 2004. Data mining in e–learning domain. *Emerald group publishing limited*, **21**(1), 29 – 34.

Hanneman, Robert A., & Riddle, Mark. 2005. *Introduction to social network methods [online text–book]*. `http://www.faculty.ucr.edu/~hanneman/`. [Accessed: February 2014].

Haythornthwaite, C. 2002. *Building social networks via computer networks: Creating and sustaining distributed learning communities*. Cambridge University Press. Pages 159 – 190.

Haythornthwaite, Caroline. 1996. Social network analysis: An approach and technique for the study of information exchange. *Library information science research*, **18**(4), 323 – 342.

Haythornthwaite, Caroline. 1998. A social network study of the growth of community among distance learners. *Information research*, **4**(1), 1 – 27.

Haythornthwaite, Caroline. 1999. Networks of information sharing among computer–supported distance learners. *Proceedings of the 1999 conference on computer support for collaborative learning CSCL 99*, 218 – 222.

Haythornthwaite, Caroline. 2001. Exploring multiplexity: Social network structures in a computer–supported distance learning class. *The information society*, **17**(3), 211 – 226.

Haythornthwaite, Caroline. 2005. Social network methods and measures for examining e–learning. *Library*, 1 – 22.

Haythornthwaite, Caroline. 2006. Learning and knowledge networks in interdisciplinary collaborations. *J. Am. Soc. Inf. Sci. Technol. JASIST*, **57**(8), 1079 – 1092.

Haythornthwaite, Caroline. 2011. Learning networks, crowds and communities. *Pages 18 – 22 of: Proceedings of the 1st international conference on learning analytics and knowledge*. LAK '11. ACM.

Hiltz, Starr Roxanne, & Turoff, Murray. 2002. What makes learning networks effective? *Communications of the ACM*, **45**(4), 56 – 59.

Huisman, Mark. 2003. Software for social network analysis. *Social science research*, **25**(2), 270 – 316.

Hunter, M. Gordon. 2004. *Qualitative research in information systems: An exploration of methods*. Idea Group Publishing. Chap. 16, pages 291 – 304.

Initiatives, EDUCAUSE Learning. 2011. 7 things you should know about MOOCs. *EDUCAUSE*, **65**, 2.

Ivanova, M. 2009. Use of start pages for building a mashup personal learning environment to support self–organized leaners. *Serdica journal of computing*, **3**, 227 – 238.

Jafari, Ali, McGee, Patricia, & Carmean, Colleen. 2006. Managing courses, defining learning: What faculty, students, and administrators want. *Educause review*, **41**(4), 50 – 71.

Johnson, L., Adams, S., & Haywood, K. 2011. The horizon report 2011. *The new media consortium.*

Jones, Chris, & Dirckinck-Holmfield, Lone. 2009. Analysing networked learning practices: An introduction. *Pages 1 – 28 of: Analysing networked learning practices in higher education and continuing professional development.* Technology Enhanced Learning, no. 4.

Jung, Insung. 2001. Building a theoretical framework of web-based instruction in the context of distance education. *British journal of educational technology*, **32**(5), 525 – 534.

Kadushin, Charles. 2005. Who benefits from network analysis: ethics of social network research. *Social networks*, **27**(2), 139 – 153.

Klamma, Ralf, Chatti, Mohamed Amine, Duval, Erik, Hummel, Hans G. K., Hvannberg, Ebba Thora, Kravcik, Milos, Law, Effie, Naeve, Ambjörn, & Scott, Peter. 2007. Social software for life–long learning. *Educational technology & society*, **10**(3), 72 – 83.

Knoke, David, & Yang, Song. 2008. *Social network analysis.* 2nd edn. Quantitative applications in the social sciences, vol. 154. SAGE.

Kop, Rita. 2011. The challenges to connectivist learning on open online networks: Learning experiences during a massive open online course. *International review of research in open and distance learning*, **12**(3), 19 – 38.

Kop, Rita, & Carroll, Fiona. 2011. Cloud computing and creativity: Learning on a massive open online course. *European journal of open, distance and e–learning*, 1 – 11.

Kop, Rita, & Hill, Adrian. 2008. Connectivism: Learning theory of the future or vestige of the past? *The international review of research in open and distance learning*, **9**(3), 1 – 13.

Krebs, V. 2004. *An introduction to social network analysis [online].* `http://www.orgnet.com/sna.html`. [Accessed: February 2014].

Krebs, V. 2006. *How to do social network analysis [online].* `http://www.orgnet.com/sna.html`. [Accessed: February 2014].

Krebs, Valdis E. 2002. Mapping networks of terrorist cells. *Connections*, **24**(3), 43 – 52.

Kumar, R., Ragbavan, P., Rajagopalan, S., & Tomkins, A. 2002. The web and social networks. *Computer*, **35**, 32 – 36.

Kylli, Jan-Peter. 2005. *The future of eLearning. a short history of eLearning and a look into the future of computer mediated learning.* `http://mlab.uiah.fi/~kavetiso/tsure/essays/eLearning.pdf`. [Accessed: February 2014].

Laghos, & Zaphiris. 2006. Sociology of student–centred e–Learning communities: A network analysis. *In: IADIS international conference.* Dublin, Ireland: e–Society.

Lee, Mark J. W., & McLoughlin, Catherine. 2007. Teaching and learning in the Web 2.0 era: Empowering students through learner–generated content. *International journal of instructional technology and distance learning*, **4**(10).

Li, Linda C., Grimshaw, Jeremy M., Nielsen, Camilla, Judd, Maria, Coyte, Peter C., & Graham, Ian D. 2009. Evolution of wenger's concept of community of practice. *Implementation science: IS*, **4**(11), 1 – 8.

Lin, Nan. 2001. *Social capital: A theory of social structure and action.* Structural analysis in the social sciences, no. 19. Cambridge University Press.

Long, Phil, & Siemens, George. 2011. Penetrating the fog: Analytics in learning and education. *Educause review*, **46**(5), 31 – 40.

manoj Parameswaran, & andrew whinston. 2007. research issues in social computing. *Journal of the association for information systems JAIS*, **8**, 336 – 350.

Marin, Alexandra, & Wellman, Barry. 2009. *Social Network Analysis: An Introduction.* Sage.

Marsden, P.V. 2002. Egocentric and sociocentric measures of network centrality. *Social networks*, **24**(4), 407 – 422.

Martínez, A., Dimitriadis, Yannis A., Rubia-Avi, Bartolomé, Gómez, E., Garrachón, I., & Marcos, José Antonio. 2002. Studying social aspects of computer–supported collaboration with a mixed evaluation approach. *Pages 631 – 632 of: CSCL.*

Martínez, A., Dimitriadis, Y., Rubia, B., Gómez, E., & de la Fuente, P. 2003. Combining qualitative evaluation and social network analysis for the study of classroom social interactions. *Computers & education*, **41**(4), 353 – 368.

Martino, Francesco, & Spoto, Andrea. 2006. Social network analysis: A brief theoretical review and further perspectives in the study of information technology. *Social networks*, **4**(1), 53 – 86.

Meyer, Katrina A. 2010. A comparison of Web 2.0 tools in a doctoral course. *The internet and higher education*, **13**(4), 226–232.

Mills, Albert J., Durepos, Gabrielle, & Wiebe, Elden. 2010. *Encyclopedia of case study research.* SAGE Publications, Inc.

Moodle, Website. 2013. *Moodle statistics*. `https://moodle.org/stats/`. [Accessed: February 2014].

Mott, Jonathan. 2010. Envisioning the Post–LMS era: The open learning network. *EDUCAUSE quarterly*, **33**(1), 1 – 8.

Müller-prothmann, Tobias. 2007. Social network analysis: a practical method to improve knowledge sharing. *Analysis*, 219 – 233.

Neal, Lisa. 2007. Predictions for 2007. *eLearn*, **2007**(1).

Newman, M. E. J. 2001a. Scientific collaboration networks. I. Network construction and fundamental results. *Physical Review E*, **64**(1).

Newman, M. E. J. 2001b. Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical Review E*, **64**(1).

Newman, M E J. 2004. Citebase – who is the best connected scientist? a study of scientific coauthorship networks. *Lecture notes in physics new york then berlin*, **650**, 337 – 370.

Niu, Huiqing. 2010. Social network analysis of university online forum. *Pages 422 – 429 of: Computational aspects of social networks (CASoN), 2010 international conference on*.

Norris, Donald M. 2010. 7 things you should know about first–generation learning analytics. *EDUCAUSE learning initiative (ELI)*, **April**.

Nurmela, Kari, Lehtinen, Erno, & Palonen, Tuire. 1999. Evaluating CSCL log files by social network analysis. *In: Proceedings of the 1999 conference on computer support for collaborative learning*. CSCL '99.

Oneto, Lucia, Herder, Eelco, Abel, Fabian, & Smits, David. 2009. Making today's learning management systems adaptive. *In: Learning in the synergy of multiple disciplines, proceedings of the EC–TEL 2009*. Lecture Notes in Computer Science, vol. 5794. Springer.

O'Reilly, Tim. 2005. *What is Web 2.0? design patterns and business models for the next generation of software*. `http://oreilly.com/web2/archive/what-is-web-20.html`. [Accessed: February 2014].

O'Reilly, Tim. 2007. What is Web 2.0: Design patterns and business models for the next generation of software. *Design*, **65**(65), 17 – 37.

Passmore, David L. 2010. *Social network analysis theory and applications*. `http://train.ed.psu.edu/WFED-543/SocNet_TheoryApp.pdf`. [Accessed: February 2014].

Pescosolido, Bernice. 2011. The sociology of social networks. *Asian journal of social psychology*, **1**, 209 – 217.

Power, Christopher, Petrie, Helen, Sakharov, Vasily, & Swallow, David. 2010. Virtual learning environments: another barrier to blended and e–learning. *Pages 519 – 526 of: Proceedings of the 12th international conference on computers helping people with special needs: Part i*. ICCHP'10. Springer-Verlag.

Rallo R., Gisbert M., & J., Salinas. 2005. Using data mining and social networks to analyze the structure and content of educative online communities. *Pages 1 – 10 of: In international conference on multimedia and ICTs in education*.

Redecker, Christine. 2009. *Review of Learning 2.0 practices: Study on the impact of Web 2.0 innovations on education and training in europe*. JRC scientific and technical reports. Institute for Prospective Technological Studies, European Commission Joint Research Centre.

Reffay, C, & Chanier, T. 2003. How social network analysis can help to measure cohesion in collaborative distance–learning. *Learning*, 1 – 10.

Reffay, Christophe, & Chanier, Thierry. 2002. Social network analysis used for modeling collaboration in distance learning groups. *Intelligent tutoring systems*, **2363**(1 September), 31 – 40.

Ressler, Steve. 2006. Social network analysis as an approach to combat terrorism: Past, present, and future research. *New york*, **2**(2), 1 – 10.

Reyes, Pablo, & Tchounikine, Pierre. 2005. Mining learning groups' activities in forum–type tools. *Pages 509 – 513 of: Proceedings of th 2005 conference on computer support for collaborative learning: learning 2005: the next 10 years!* CSCL '05.

Rogers, E.M. 2003. *Diffusion of innovations*. 5th edn. New York, Free press.

Rogers, Jim. 2000. Communities of practice: A framework for fostering coherence in virtual learning communities. *Educational technology & society*, **3**(3), 384 – 392.

Romero, C, & Ventura, S. 2007. Educational data mining: A survey from 1995 to 2005. *Expert systems with applications*, **33**(1), 135 – 146.

Romero, C, & Ventura, S. 2010. Educational data mining: A review of the state of the art. *IEEE transactions on systems man and cybernetics part c applications and reviews*, **40**(6), 601 – 618.

Romero, C, Ventura, S, & Garcia, E. 2008. Data mining in course management systems: Moodle case study and tutorial. *Computers & education*, **51**(1), 368 – 384.

Rovai, Alfred P. 2002. Sense of community, perceived cognitive learning, and persistence in asynchronous learning networks. *The internet and higher education*, **5**(4), 319 – 332.

Safran, Christian, Helic, Denis, & Guetl, Christian. 2007. E-learning practices and Web 2.0. *Computer*, **1**(1), 1 – 8.

Saltz, Jeffrey S, Hiltz, Starr Roxanne, & Turoff, Murray. 2004. Student social graphs: Visualizing a student's online social network. *Networks*, 596 – 599.

Schmees, Markus. 2006. Organizing technology enhanced learning. *Pages 139 – 150 of: Proceedings of the 8th international conference on electronic commerce: The new e-commerce: innovations for conquering current barriers, obstacles and limitations to conducting successful business on the internet*. ICEC '06. ACM.

Scott, John. 2000. *Social network analysis: A handbook*. 2nd edn. SAGE.

Shane Dawson, Erica McWilliam, & Poole, Gary. 2008. Monitoring student creative capacity: using network visualization to evaluate pedagogical practice. *Centre of excellence for creative industries and innovation (CCI)*. [Accessed: February 2014].

Shen, D., Nuankhieo, P., Huang, X., Amelung, C., & Laffey, J. 2008. Using social network analysis to understand sense of community in an online learning environment. *Journal of educational computing research*, **39**(1), 17 – 36.

Sie, Rory, Ullmann, Thomas, Rajagopal, Kamakshi, Cela, Karina, Bitter-Rijpkema, Marlies, & Sloep, Peter. 2012. Social network analysis for technology–enhanced learning: review and future directions. *J–JUCS*, **18**(1), 44 – 61. [Accessed: February 2014].

Siemens, G. 2010. *What are learning analytics?* `http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/`. [Accessed: February 2014].

Siemens, G., & Downes, S. 2008. *Connectivism & connected knowledge*. `http://ltc.umanitoba.ca/connectivism/`. [Accessed: February 2014].

Siemens, G., Downes, S., Cormier, D., & Kop, R. 2010. *PLENK2010 – personal learning environments, networks and knowledge*. `http://connect.downes.ca/`. Course Blog. [Accessed: February 2014].

Siemens, George. 2004. *Learning management system: The wrong place to start learning*. `http://www.elearnspace.org/Articles/lms.htm`. [Accessed: February 2014].

Siemens, George. 2005a. Connectivism: A learning theory for the digital age. *In: International journal of instructional technology and distance learning*.

Siemens, George. 2005b. *Learning development cycle: Bridging learning design and modern knowledge needs*. `http://www.elearnspace.org/Articles/ldc.htm`. [Accessed: February 2014].

Siemens, George. 2006. *Knowing knowledge*. Lulu.com. [Accessed: February 2014].

Soy, Susan K. 1997. *The case study as a research method*. `https://www.gslis.utexas.edu/~ssoy/usesusers/l391d1b.htm`. [Accessed: February 2014].

Steinert, Anne. 2010. ConnectLearning – an answer for the new challenges? *eLearning papers*, **18**(February), 1 – 13.

Tery Martindale, Michael Dowdy. 2010. Personal learning environments. *In: Emerging technologies in distance educatioin.*

Tim Berners-Lee, James Hendler, & Lassila, Ora. May 2001. The semantic web. *Scientific american*, 29 – 37.

Travers, Jeffrey, & Milgram, Stanley. 1969. An experimental study of the small world problem. *Sociometry*, **32**(4), 425 – 443.

Valiathan, Purnima. 2002. *Blended learning models.* `http://www.learningcircuits.org/2002/aug2002/valiathan.html`. [Accessed: February 2014].

van Harmelen, Mark. 2006. Personal learning environment. *Pages 815 – 816 of: ICALT '06 proceedings of the sixth IEEE international conference on advanced learning technologies.*

Vassileva, Julita. 2008. Toward social learning environments. *IEEE transactions on learning technologies*, **1**(4), 199 – 214.

Vázquez, Alexei. 2001 (May). *Statistics of citation networks.* `http://arxiv.org/abs/cond-mat/0105031`. [Accessed: February 2014].

Vygotsky, L.S. 1978. *Mind in society: The development of higher psychological processes.* Harvard University Press.

Wasserman, S., & Galaskiewicz, J. 1994. *Advances in social network analysis: Research in the social and behavioral sciences.* Sage, Thousand Oaks, CA.

Wasserman, Stanley, & Faust, Katherine. 1994. *Social network analysis: Methods and applications.* Structural analysis in the social sciences, vol. 8. Cambridge university press.

Watts, D. J., & Strogatz, S. H. 1998. Collective dynamics of "small–world" networks. *Nature*, **393**(6684), 440 – 442.

Watts, Duncan J. 2004. The "new" science of networks. *Annual review of sociology*, **30**, 243 – 270.

Weaver, Alfred C., & Morrison, Benjamin B. 2008. Social networking. *IEEE computer*, **41**(2), 97 – 100.

Wellman, B. 2001. Computer networks as social networks. *Science*, **293**, 2031 – 2034.

Wellman, Barry. 1988. Structural analysis: from method and metaphor to theory and substance. *Pages 19 – 61 of: Social structures: a network approach.* Cambridge University Press.

Wellman, Barry. 1996. *For a social network analysis of computer networks: a sociological perspective on collaborative work and virtual community.* ACM Press. Pages 1 – 11.

Wenger, Etienne. 1998. *Communities of practice: learning, meaning, and identity.* 1 edn. Cambridge University Press.

Wetherell, Charles. 1998. Historical social network analysis. *International review of social history*, **43**, 125 – 144.

Willging, P. A. 2005. Using social network analysis technique to examine online interaction. *US–China education review*, **2**(9), 46 – 56.

William, Rice. 2006. *Moodle e–learning course development: A complete guide to successful learning using Moodle.* 2nd edn. SAGE.

Williams, J.B., & Jacobs, J. 2004. Exploring the use of blogs as learning spaces in the higher education sector. *Australasian journal of educational technology*, **20**(2), 232 – 247.

Wilson, Scott. 2005. *The future of VLE.* `http://zope.cetis.ac.uk/members/scott/blogview?entry=20050125170206`. [Accessed: February 2014].

Wilson, Scott, Liber, Oleg, Johnson, Mark William, Beauvoir, Phillip, Sharples, Paul, & Milligan, Colin D. 2006. Personal learning environments: Challenging the dominant design of educational systems. *In: EC-TEL workshops.*

Yang, Heng-Li, & Tang, Jih-Hsin. 2004. Team structure and team performance in IS development: a social network perspective. *Information & management*, **41**(3), 335 – 349.

Yin, R. K. 1984. *Case study research: Design and methods.* Newbury Park, CA: Sage.

Zaiane, Osmar R. 2001. Web usage mining for a better web–based learning environment. *Advances*, 60 – 64.

Zaphiris, Panayiotis, Laghos, Andrew, & Zacharia, Giorgos. 2009. Distributed construction through participatory design. *Pages 1181 – 1185 of: Encyclopedia of information science and technology (II).* IGI Global.

Zhang, Lei, & Tu, Wanqing. 2006. Six degrees of separation in online society. *Distribution*, **3**(12), 1 – 5.