MINISTRY FORDEVELOPMENT OFINFORMATION TECHNOLOGIES ANDCOMMUNICATIONSOF THE REPUBLIC OFUZBEKISTAN

TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES

On the manuscript right

UDC004.6

RUSTAMOVA ZULFIYA DILSHODOVNA

Analysis and development of transcended scientific knowledge base based on social networking

Specialization: 5A330501 – Computer engineering
(Applied software development)

Dissertationfor getting master degree

The work considered and allowed
for the defense
Ph.D., head of the department of
Multimedia Technologies
Nazirova E.

____

«____»___________________2015y.

Supervisor
Ph.D., Associate Professor
Sadullayeva SH. A.

____

«____»_ __________________2015y.

Tashkent-2015

Content

Introduction

Subject topicality. Uzbekistan has been struggling to bring its telecom system up to the standard found in developed countries. Although steadily improving, the telecommunications infrastructure remains outmoded and inadequate. However, since 2002 the situation has been gradually improving, due largely to the government's decision to priorities ICTs. Consequently there has been an upward trend in the country's telecom market over recent years, with rising revenues and increased investment in infrastructure. Government strategic policy is to privatize the incumbent operator Uzbektelecom and to open the market to competition in accordance with the country's aim to join the World Trade Organization (WTO).Uzbek internet users are enjoying cheaper and faster access than they have ever known. Starting August 1, the maximum internet speed in Uzbekistan reached 11.8 Gbps. "The process is rapidly unfolding," Sherzod Shermatov, deputy chairman of the State Committee for Communication, Informatization and Telecommunication Technologies, said. "During the past six months, speeds have increased by a whole 15%."

Last decades foreign languages are object of steadfast attention and studying not only scientific, but also simple people. The increasing role, which is played by foreign languages, in realization of influence on consciousness and activity of various strata of society is thus marked. Also it is necessary to consider that the knowledge of foreign languages is rather significant in sphere of personal and professional communications of the person that allows him to be to on a step ahead of others. The importance of knowledge of a foreign language (more often English) is difficult for overestimating. It is almost impossible to imagine life of the modern person who does not know a foreign language, after all the majority of modern communication media and dialogue is focused on people to some extent, who knows the language.

To execute the decree of the President of the Republic of Uzbekistan from December, 10th, 2012 № «About measures on the further perfection of system learning foreign languages» PP-1875: to confirm the State educational standard of system of continuous formation «Requirements to level of readiness of graduates of all steps of formation on foreign languages». To the ministry of the higher and secondary special education, to the State testing center under the Cabinet of Ministers of the Republic of Uzbekistan, to the Tashkent university of information technologies [1]:

- since 2013/2014 academic years to include in test tests at reception in the Tashkent university of information technologies in all directions of formation the block of English language as an obligatory subject, and also to enter, since 3rd course of a Bachelor Degree and a Magistracy teaching of separate profile and special subjects in English domestic and involved foreign teachers;

- from 2015/2016 academic year to enter practice in separate groups of a Bachelor Degree and a Magistracy studying common professional and special disciplines in English with the use of profile foreign language manuals and literature.

Internet is a very essential part of life from shopping to electronic mails and education, internet is very important. It is a very large community which is using internet for pure education but unfortunately we have also a very large number of people including majority of youth and teenager using Internet only for social networks. Internet is very big evolution of technology but when we talk about the social networks it is extremely dangerous for youth and become enormously common and widespread in past few years.

Social media and knowledge management may seem to be the same thing based on their basic characteristics, but in reality they are different. I am not going to argue that one is better than the other, playing that type of zero-sum game is a waste of time. Rather than a ‘these’ vs. ‘those’ argument, it is time to recognize the differences and move on to figure out how best to apply each.

Equating social media to knowledge management makes sense if there is only one way to create, serve, and consume knowledge. Thankfully there are many ways and that makes social media different from knowledge management.

Object and subject of research. Object of research is educational social networking. Subject of research is analysis and development of transcended scientific knowledge base based on social networking using semantic web and also expert system.

Research purpose. The purpose of work consists of the research on existing major analyzing of social networking knowledge bases with are based on social media, conduction of a comparative analysis of these algorithms, development and implementation of a unique scientific knowledge base using program language PHP.

Research objectives:

– review of the sources of literature devoted to the solution of educational social networking problems;

– development of a social knowledge base unique parameter for existing social networking;

– implementation of social media computer software in the field of programming PHP using frames based on unique proposed parameter;

– testing of the developed computer software in the real-life situations.

Research methods. To achieve the goal were used methods of empirical research, probability theory, algorithms and methods of developing educational social networking base.

Scientific novelty. Scientific novelty of research work is a scientific network, developed on the expert system ability to unite scientists on their specialized fields, topics, interests, etc. This system includes as a base for storing scientific publications and on the basis of academics, professors (scientific figures).

Practical value. Purposed software on this research work by internet by connecting students and tutors, academics and other teachers. This connection will be only by science. User be able to register in, publish their topics, papers, and other scientific publications.

Publications. On the subject of dissertation following 2 articles have been published:

– “U-healthcare efficient process and data management system” in the International conference “Perspectives for the development of information technologies ITPA 2014” held in November 4-5, 2014 in TUIT.

– “Role-Based Access Control for SCADA Security access control” in scientific journal named “TATU XABARLARI” on September, 2015 in TUIT.

Structure and content. Master’s dissertation work consists of introduction, three chapters, conclusion, list of used references and application.

Dissertation structure. There is introduction, three chapters, conclusion, references and attachment.

First chapter devotes to main concepts, major terms, benefits and structure of decision support system,social networking, tradition education and advantages -disadvantages of social networking, problems ofa knowledge base is in data quality,a conception of the semantic web expert system.

The second chapter types of structure of social networking, algorithms and metaheuristic algorithms social network and semantic web task solving, development of social network analysis, algorithms of social network and semantic web,social networking system algorithm,semantic matchmaking algorithm

Final chapter consist mainly information about software, its structure, modules and user manual.

In conclusion there is general concludes which comes from dissertation.

ChapterI. Basic knowledge of social networking, knowledge bases and semantic web

1.1. Social networking, tradition education and advantages - disadvantages of Social Networking

Social networking is the grouping of individuals into specific groups, like small rural communities or a neighborhood subdivision, if you will. Although social networking is possible in person, especially in the workplace, universities, and high schools, it is most popular online.

This is because unlike most high schools, colleges, or workplaces, the internet is filled with millions of individuals who are looking to meet other people, to gather and share first-hand information and experiences about cooking, golfing, gardening, developing friendships professional alliances, finding employment, business-to-business marketing and even groups sharing information about baking cookies to the Thrive Movement. The topics and interests are as varied and rich as the story of our universe.

When it comes to online social networking, websites are commonly used. These websites are known as social sites. Social networking websites function like an online community of internet users. Depending on the website in question, many of these online community members share common interests in hobbies, religion, politics and alternative lifestyles. Once you are granted access to a social networking website you can begin to socialize. This socialization may include reading the profile pages of other members and possibly even contacting them.

The friends that you can make are just one of the many benefits to social networking online. Another one of those benefits includes diversity because the internet gives individuals from all around the world access to social networking sites. This means that although you are in the United States, you could develop an online friendship with someone in Denmark or India. Not only will you make new friends, but you just might learn a thing or two about new cultures or new languages and learning is always a good thing.

As mentioned, social networking often involves grouping specific individuals or organizations together. While there are a number of social networking websites that focus on particular interests, there are others that do not. The websites without a main focus are often referred to as "traditional" social networking websites and usually have open memberships. This means that anyone can become a member, no matter what their hobbies, beliefs, or views are. However, once you are inside this online community, you can begin to create your own network of friends and eliminate members that do not share common interests or goals.

As I'm sure you're aware, there are dangers associated with social networking including data theft and viruses, which are on the rise. The most prevalent danger though often involves online predators or individuals who claim to be someone that they are not. Although danger does exist with networking online, it also exists in the real world, too. Just like you're advised when meeting strangers at clubs and bars, school, or work - you are also advised to proceed with caution online.

By being aware of your cyber-surroundings and who you are talking to, you should be able to safely enjoy social networking online. It will take many phone conversations to get to know someone, but you really won't be able to make a clear judgment until you can meet each other in person. Just use common sense and listen to your inner voice; it will tell you when something doesn't feel right about the online conversations taking place. Once you are well informed and comfortable with your findings, you can begin your search from hundreds of networking communities to join. This can easily be done by performing a standard internet search. Your search will likely return a number of results, including MySpace, FriendWise, FriendFinder, Yahoo! 360, Facebook, Orkut, and Classmates.

What are Social Network Sites and How to Use Them? When a teenager joins a site like Facebook they first create a personal profile. These profiles display information such as one's name, relationship status, occupation, photos, videos, religion, ethnicity, and personal interests. What differentiates SNS from previous media like a personal homepage is the display of one's friends. In addition to exhibiting a network of friends, other users can then click on their profiles and traverse ever widening social networks. These three features—profiles, friends, traversing friend lists—represent the core, defining characteristics of SNS. Social networking features are increasingly integrated into other types of media tools and online communities. Sonia Livingstone notes that SNS invite "convergence among the hitherto separate activities of email, messaging, website creation, diaries, photo albums and music or video uploading and downloading" (p. 394). For example, YouTube is primarily a video sharing service, but users can add others as their friends or subscribe to a member's collection of videos. Using boyd and Ellison's definition, YouTube can be included as a type of SNS. As researchers examine the effects of SNS on social behaviors, they will undoubtedly come across these blurring of technologies. The proliferation of SNS, both as standalone communities and integrated into other media tools, underscores the importance of understanding the unique effects these sites have on human interaction. Amid the sea of what websites can be termed SNS, the technical definition of SNS still provides a shared conceptual foundation. Comparing across common features—i.e., profiles and friend networks—researchers can begin to understand how various communities co-opt these characteristics to create entirely new cultural and social uses of the technology. Lange's ethnographic study of YouTube shows that users deal with issues concerning public and private sharing of video. Some YouTube users post videos intended for wide audiences, but share very little about their own identities. Their motivations might be to achieve Internet fame and gather viewers. Other members upload videos intended for a small network of friends and may restrict the privacy settings to only allow access to those individuals. The concepts of friend and social network for these users are entirely distinct. The features and culture of particular SNS communities may also affect behavior. Pappacharissi analyzes profiles and user behavior on Facebook, LinkedIn, and ASmallWorld and finds that the features, intent, and norms of each social network are intricately related to user behavior. For example, Facebook is a more wide-open network with less stringent rules about membership, information disclosure, and interaction. LinkedIn is also an open-membership network, but its design such as profiles in résumé-like format encourages professional uses. A Small World is an entirely closed and exclusive network, where members share pictures that signal their socioeconomic status. Similarly, when MySpace introduced its Top 8 function, where users designated their top friends on their profile, it set off a firestorm of social drama among teens. "There are tremendous politics behind the Top 8, not unlike the drama over best and best friends in middle school" (para. 32). These examples highlight how the structure, function, and mission of a respective SNS community influence networking behavior. Signaling Theory, Warranting Theory, and Identity Development The process of creating profiles has been a major focus of theoretical and empirical discussion. The common features of profiles include personal information such as one's name, location, school affiliation, occupation, and personal interests such as favorite movies or music. Other vital components of the profile are pictures, videos, and the comments one's peers leave on the page. Profiles can be updated at any time and some sites like MySpace allow individuals control as to how their profile looks. Using programming techniques, youth frequently apply "skins" to their MySpace profiles that completely alter the visual design or interface of their pages. Signaling theory is one framework used to understand how individuals disclose information on their SNS profiles. Donath observes that, "Whether face-to-face or online, much of what people want to know is not directly observable" (para. 10). She contends that much of human interaction consists of signals that communicate the status and characteristics of an individual. Signaling theory examines how one's self-presentation in SNS develops identity and trust with others. For example, when a user displays a contact as a "friend" he or she is—in an indirect way—vetting that that person is in fact who they claim to be. Thus, members who indiscriminately add any and all friend requests (including fake profiles or people they do not know) in an effort to seem popular may instead damage their credibility and trustworthiness to others. Among teenagers, boyd finds that "it is cool to have Friends on MySpace but if you have too many Friends, you are seen as a MySpace whore" (p. 129). In a similar vein, warranting theory suggests that human beings do in fact judge others based on cues in SNS profiles. Walther and colleagues have shown that an individual is consistently rated as physically and socially attractive when his or her friends are also attractive. Positive and negative comments left on a person's Facebook wall also greatly influence whether they are seen as attractive. In addition to judging others based on their profiles, SNS users appear to judge the credibility of profile information quite consistently. On SNS we are judged by the company we keep. Signaling theory and warranting theory also propose that people assess other-generated statements as more credible compared to self-generated information. This hypothesis is especially likely in SNS because profile owners can manipulate what information is presented on their page. Thus, statements from others might be seen as more credible than statements from the individual. Early experiments show that a Facebook user is rated as more attractive if others state that identity (through wall posts, comments, etc.) compared to when the individual (through self-statements on the profile) asserts this identity SNS profiles not only represent information that an individual chooses to disclose, but also signal what those friends indicate about the individual. These early studies offer compelling evidence that what one puts on one's SNS profile is assessed by others and the characteristics of friends are strongly related to how one is viewed. In addition, the feedback provided by one's network in an SNS is influential in the development of social identity.

Adolescents use SNS in a variety of ways. They disclose personal information about their identities and tastes on their profiles. Teenagers must also add or reject friend requests from their peers, navigating the complicated web of friendship practices. Finally, the interactions and feedback that one's network provides in SNS —through wall posts and comments—show how complex social identity and peer influence processes occur in these online communities. Social network sites provide a platform for teenagers to develop personal and social identities. Developing identities in SNS is very similar to offline contexts. Donath and boyd observe some of the ways that individuals reflect their social identity: In the physical world, people display their connections in many ways. They have parties in which they introduce friends who they think would like—or impress—each other. They drop the names of high status acquaintances casually in their conversation. They decorate their refrigerator with photos. Simply appearing in public with one's acquaintances is a display of connection. In similar fashion, young adult users of SNS decide what to place on their profiles and what friends to display for others to see. The emerging picture is that individuals make explicit decisions to disclose information about themselves on their profiles, and their networks provide social feedback to those profile displays. This process of developing identity is quite salient to adolescents who are experiencing a time of rapid growth and development. The majority of current research on SNS attempts to understand the phenomena itself. Scholars have been interested in how youth use these technologies, what cultural practices emerge in these online contexts, and what theoretical implications SNS have on personal identity and social relationships. The early descriptive and ethnographic research on youth, Internet, and social media offer rich evidence that (a) the features of different platforms, for example, the MySpace Top 8 case, influence the social practices of youth within those online communities, (b) SNS are important places for youth to develop their personal identity, and (c) youth use technologies like SNS to mediate their relationships with friends, romantic partners, and broader groups of peers. The questions that parents, educators, and researchers now grapple with concern the effects SNS have on adolescent outcomes.

The use of distance education courses as a primary instructional delivery option, especially in the higher education community, is expanding at an unprecedented rate. The 9.7% growth rate in the number of college and university students enrolled in at least one online class reported by Allen and Seaman significantly exceeded the 1.5% growth rate in the overall higher education student population during the same period. Simultaneously, the emergence and growth of commercial social networking sites such as Facebook, Friendster, LinkedIn, LiveJournal, and MySpace has been extensive and widespread. Facebook, for example, is currently the fastest growing commercial SNS in the world, with more than 300 million active user profiles. Given the rising popularity of both distance education and SNSs, it seems logical to merge these popular two technologies with the goal of improving online teaching and learning.

Distance education courses are often more successful when they develop communities of practice as well as encourage high levels of online social presence among students. Fostering a sense of community is critically important, especially in an online environment where students often do not get the opportunity to meet face-to-face with other students or the instructor in the course. Since they facilitate the sharing of information—personal and otherwise—the technologies used in SNSs aid discussion and create intimacy among online students, as they have the ability to connect and build community in a socially and educationally constructed network.

In contrast to SNS, course management systems, such as Blackboard and Moodle, tend to be very focused and lack the personal touch and networking capacity that SNSs offer. For example, instructors using CMS may pose a question in an online discussion board and each student posts a response. However, these student posts are really not interactions at all, but merely question and answer sessions. Using an SNS that is user centered, rather than class centered, such as a CMS, has the potential to increase student engagement. SNSs can actively encourage online community building, extending learning beyond the boundaries of the classroom. A comparison of typical SNS and a traditional CMS appears in Table 1.

Table 1.1.

A comparison of typical SNS and a traditional CMS

Comparison of SNS and CMS Tools		
Tools	SNS	Traditional CMS
Forum	X	X
Blog	X	X
Media Sharing	X	
Messaging	X	X
Wiki		
RSS	X	
Chat	X	X
Calendar	X	X
Tagging	X	
Own Brand & Visual Design	X	
Real time Activity Stream	X	
Groups	X	
Friends	X	
Profile Pages	X	
File sharing		X

While commercial SNSs, such as Facebook and MySpace, are popular, newly emerging SNSs created specifically for an educational audience provide a unique opportunity for educators to “facilitate a strong sense of community among students” and encourage “personal interactions that can lead to the creation of new knowledge and collective intelligence”. In order to evaluate the largely unexplored educational benefits of SNSs, we surveyed graduate students enrolled in distance education courses using Ning in Education, an education-based SNS, based on their attitudes toward SNSs as productive online tools for teaching and learning. Our results suggest that education-based SNSs, such as Ning in Education can be used most effectively in distance education courses as a technological tool for improved communication among students at the higher education level.

Analysis of social network diagrams helps determine the extent to which certain people are central to the effective functioning of a network, regardless of whether or not divisive subgroups in a network exist or what the overall connection of a given network is. Things to look for in SNA:

- Bottlenecks—Central nodes that provide the only connection between different parts of the network.
- Number of links—Insufficient or excessive links between departments that must coordinate effectively.
- Average distance—Degrees of separation connecting all pairs of nodes in the group. Short distances transmit information accurately and in a timely way, while long distances transmit slowly and can distort the information.
- Isolation—People that are not integrated well into a group and therefore, represent both untapped skills and a high likelihood of turnover.
- Highly expert people—Not being utilized appropriately.
- Organizational subgroups or cliques—Can develop their own subcultures and negative attitudes toward other groups.

Five Advantages of Social Networking1. Worldwide Connectivity. No matter if you are searching for that former college roommate, your first grade teacher, or an international friend, there is no easier or faster way to make a connection than via the social network. Although Facebook, Twitter, LinkedIn and MySpace are probably the most well known social networking communities, there are new websites popping up regularly that are dedicated to allowing people to connect and to interact via the Internet. Through such sites, individuals make new friends or business connections or extend their personal base by connecting and interacting with friends of friends and so forth. These connections can help one with a variety of things such as:

- Finding romance
- Seeking a new job
- Locating assistance
- Getting and giving product and service referrals
- Receiving support from like-minded individuals
- Making or receiving advice on career or personal issues

In many ways, social communities are the virtual equivalent of meeting at the general store or at church socials to exchange news and get updated on friends and families. Snail mail pen pals have been replaced by virtual avatars and private messages sent via the social network.

2. Commonality of Interest. When you opt to participate in a social network community, you can pick and choose those individuals whose likes and dislikes are similar to yours and build your network around those commonalities. For instance, if you are a chess aficionado or a book lover, you can find and interact with those who share your interest. Because you are connecting digitally instead of having to physically attend meetings, you have the luxury of joining many more groups and communities. You can meet with your friends anytime you have an Internet connection and whenever you find them online.

3. Real-Time Information Sharing. Many social networking sites incorporate an instant messaging feature, which means you can exchange information in real-time via a chat. This is a great feature for teachers to use to facilitate classroom discussions. A study by the John D. and Catherine T. MacArthur Foundation shows these networks can be used as effective vehicles for students to pursue self-paced online learning. In addition, the Internet is the

ultimate online textbook. Students no longer need to take out six library books at a time. Much of what they need to know they can find online. School is not the only setting where this type of real-time information sharing can be beneficial. Social networking can provide a tool for managers to utilize in team meetings, for conference organizers to use to update attendees and for business people to use as a means of interacting with clients or prospects.

4. Free Advertising. Whether you are non-profit organization who needs to get the word out about your upcoming fundraiser or a business owner marketing a new product or service, there's no better way to get your message in front of millions of people 24/7. The best part is it that you can spread the word through social networking profiles for free. You can promote one product, service or idea or many because you are limited only by the amount of time you wish to invest in the endeavor.

5. Increased News Cycle Speed. Social networking has revolutionized the speed of the news cycle. Many news organizations now partner with social networking sites like Twitter, YouTube, and Facebook in order to both collect and share information. One can get a sense of what is going on in the world just by watching trending topics from many of these sites. This has led to the development of a near instantaneous news cycle as millions of social networking updates rapidly spread news and information.

Five Disadvantages of Online Social Communities.Face to Face Connections are Endangered. A huge advantage of these social communities has a reverse side effect that is also a big disadvantage of social networking: they reduce or eliminate face-to-face socialization. Because of the autonomy afforded by the virtual world, individuals are free to create a fantasy persona and can pretend to be someone else.

It is hard to say no, be rude, or ignore someone when you are looking them in the eye. It's incredibly easy and quick to unfriend or unfollow someone or simply block their efforts to make a connection. Just one click of the mouse and your problems are over. Unfortunately, this feature of online socialization cheats people of the opportunity to learn how to resolve conflicts in the world outside the Internet and it could retard or cripple one's social skills developments.

Tweens and teens are at higher risk because those years are when they are learning to interact with others or build and maintain relationships. A report from the National School Boards Association shows that of the children in these age groups that use a social network, 41 percent spend their time posting messages. They are not spending this time in face-to-face interactions with their peers or others nor are they developing the necessary social skills for future success.

2. Cyberbullying and Crimes Against Children. Use of social networks can expose individuals to harassment or inappropriate contact from others. Unless parents are diligent to filter the Internet content to which their families are exposed, children could be exposed to pornography or other inappropriate content. The Pew Center, in their Cyberbullying 2010 report, states that 93 percent of teens aged 12 to 17 use the Internet. Of that 93 percent, 63 percent of them use the Internet daily. Such high usage increases the risk of their being victims of cyberbullying or other cybercrimes.

3. Risks of Fraud or Identity Theft. Whether you like it or not, the information you post on the Internet is available to almost anyone who is clever enough to access it. Most thieves need just a few vital pieces of personal information to make your life a nightmare and if they successfully steal your identity, it could cost you dearly. A report on CNET reveals over 24 million Americans put their personal information at risk by posting it on public sites such as social communities.

4. Time Waster. A Nielsen report explains that social networking can be a big waste of time that sucks 17 percent of our Internet time down the non-productivity drain. While it is true that some of that time is likely spent in making and maintaining important business, social or professional connections, it is also true that it is easy to become distracted and end up spending valuable time on games, chats or other non-related activities. Dorie Clark of the Huffington Post reports Facebook users spend about six hours each month on the site, while social networkers spend three times as much time on those communities as they do on other online activities like email.

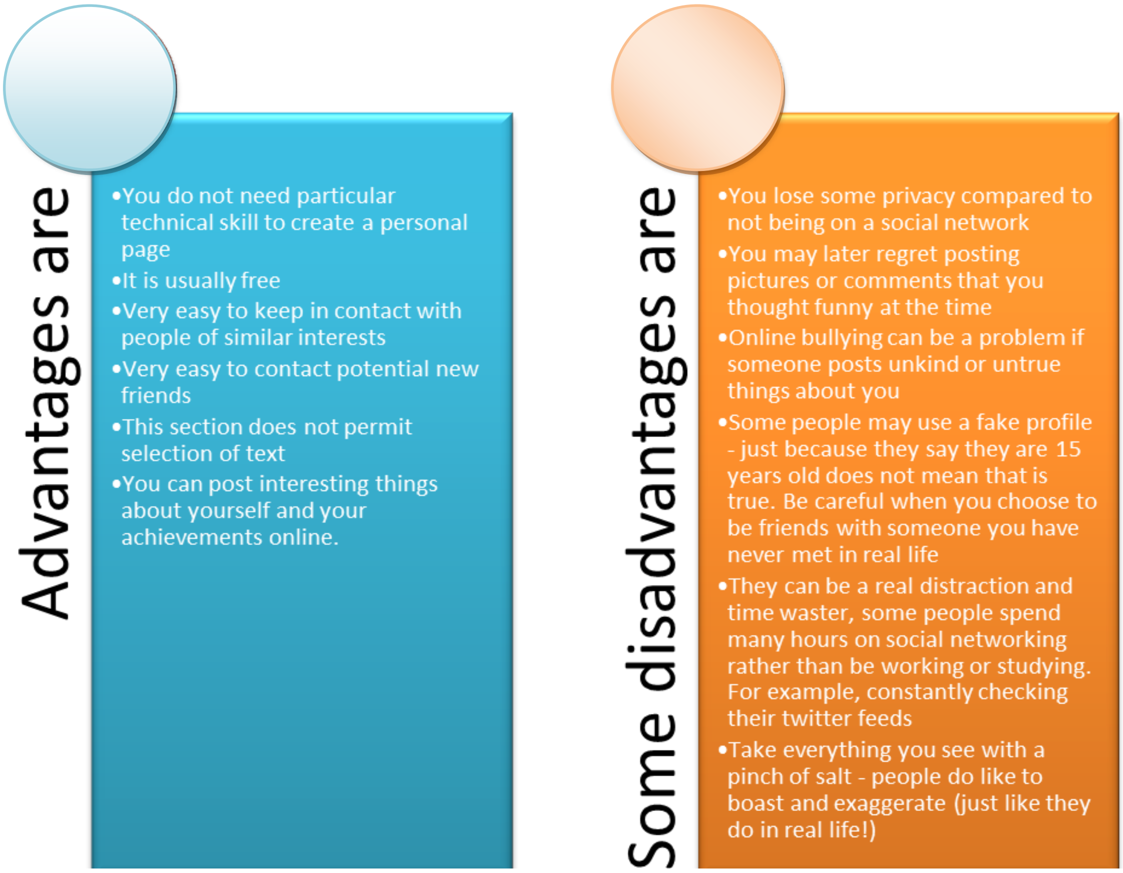


Figure1.1. Advantages and disadvantages of social education

5. Corporate Invasion of Privacy. Social networking invites major corporations to invade your privacy and sell your personal information. Have you ever posted a comment on Facebook, only to notice an advertisement appear with content related to your post? Facebook projects it will earn $3.8 billion in revenue in 2011. That's not bad for a free site. If Facebook and other social networking sites don't charge their members, however, how do they make so

much money? They do it by selling the ability to specifically target advertisements. On social networking sites, the website isn't the product, its users are. These sites run algorithms that search for keywords, web browsing habits, and other data stored on your computer or social networking profile and provide you with advertisements targeted specifically to you. At the same time, you may be giving the site permission to share your information with outside sources unless you specifically generate settings that disallow them to do so. Participating in applications like Farmville may also be allowing outside vendors access to your private information.

Take responsibility for your own safety and integrity and never join something just because it is trendy or all your friends are doing it. In evaluating the advantages and disadvantages of social networking, it's best to err on the side of caution and information. After all, the lack of both can have a devastating effect.

Every site allows you to set privacy settings. These can be changed from their default settings to limit what other people can see and read about you. For example, you could set your pages to be only viewed by friends. Other parts could be made public. Other parts could be set to family only.

1.2. Basic problems A knowledge base is in Data Quality

Social media and knowledge management may seem to be the same thing based on their basic characteristics, but in reality they are different. Equating social media to knowledge management makes sense if there is only one way to create, serve, and consume knowledge.

The descriptions may sound harsh and based in favor of social media and to some extent that is true. Knowledge should be like water — free flowing and permeating down and across your organization filling the cracks, floating good ideas to the top, lifting everyone in the organization.

Knowledge management, in practice, reflects a hierarchical view of knowledge to match the hierarchical view of the organization. Knowledge may originate anywhere in the organization, but under knowledge management it is channeled and gathered together in a knowledge base (cistern) where it is distributed based on a predefined set of channels, processes and protocols.

Social media looks chaotic in comparison. There is no predefined index, now prequalified knowledge creators, no knowledge managers, ostensibly little to no structure. Where an organization has a roof, gutters and cistern to capture knowledge, a social media organization has no roof allowing the rain to fall directly into the house collecting in puddles wherever they happen to form. That can be quite messy and organizations abhor a mess.

It is no wonder that executives, knowledge managers and software companies seek of offer tools, processes and approaches to 'tame' the social media. After all we cannot have employees, customers, suppliers and anyone else creating their own information, forming their own opinion and expressing that without our say. Think of the impact on our brand, our people, our customers… We need to manage this. We need knowledge management.

This is exactly the wrong attitude for one simple reason. It does not stop people from talking about you. Your people, customers, suppliers, competitors etc. will talk about you whenever, wherever and however they want. Sure in the past these conversations were not readily available across the World Wide Web, but they were happening. But now is not the time to seek control as much as its time to engage everyone.

Leaders recognize that engagement is the best way to glean value from the knowledge exchanged in social media. The do this, not by seeking to control social media with traditional knowledge management techniques. That only leads to what Anthony Bradley calls a 'provide and pray' approach. Translating your Lotus Notes Databases or Corporate Intranet is not the answer. It only swaps out technology without recognizing the innate difference between social media and knowledge management. If your KM capabilities were poor, adding social media will lead to the same old result only on new and different technology.

If social media is not knowledge management, then you need a different approach to create value out of social media — you need to become a social organization. Answering the question of, how do organizations gain value from social media, particularly in situations here they have not been successful with knowledge management rests in a new view of collaboration — mass collaboration. Mass collaboration consists of three things: social media, a compelling purpose and a focus on forming communities

- Social media technology provides the conduit and means for people to share their knowledge, insight and experience on their terms. It also provides a way for me to see and evaluate that knowledge based on the judgment of others.

- Purpose is the reason why people participate their ideas, experience and knowledge. They participate personally in social media because the value and identify with the purpose. They do so because they want to, rather than being told to as part of their job.

- Communities are self-forming in social media. Communities in knowledge management are often assigned by job classification or 'encouraged' based on work duties. Participation becomes prescribed creating the type of 'mandatory fun' that is the butt of Dilbert cartoons and TV sit coms. Knowledge management assigns communities because it sees knowledge as a hierarchy. Social media allows them to emerge as a property of the purpose and the participation using the tools. This lack of structure creates the space for active and innovative communities.

Making these factors work and create mass collaboration involves more than building technology and telling people to participate. It involves a range of vision, strategy and management actions that we will discuss in subsequent blog posts.

The point here is that while they may seem similar, knowledge management and social media are not the same. Recognizing the differences is a crucial step to getting value out of both and avoiding a struggle of one over the other.

Knowledge bases. The analysis of the importance of different types of regional innovation systems must take place within a context of the actual knowledge base of various industries in the economy, as the innovation processes of firms are strongly shaped by their specific knowledge base. In this paper, we shall distinguish between two types of knowledge base: analytical and synthetic. These types indicate different mixes of tacit and codified knowledge, codification possibilities and limits, qualifications and skills, required organizations and institutions involved, as well as specific competitive challenges from a globalizing economy, which have different implications for different sectors of industry, and, thus, for the kind of innovation support needed. The traditional constellation of industrial clusters surrounded by innovation supporting organizations, constituting a regional innovation system, is nearly always to be found in contexts of industries with a synthetic knowledge base (e.g. engineering-based industries), while the existence of regional innovation systems as an integral part of a cluster will normally be the case of industries-based on an analytical knowledge base (e.g. science-based industries, such as IT and bio-tech). In the discussion of different types of regional innovation systems five empirical illustrations from a Nordic comparative project on SMEs and regional innovation systems will be used: the furniture industry in Salling, Denmark; the wireless communication industry in North Jutland, Denmark; the functional food industry in Scania, Sweden; the food industry in Rogaland, Norway and the electronics industry in Horten, Norway. We argue that in terms of innovation policy the regional level often provides a grounded approach embedded in networks of actors acknowledging the importance of the knowledge base of an industry.

Knowledge Bases and Domains. A knowledge base is in Data Quality Services (DQS). To cleanse data, you have to have knowledge about the data. To prepare knowledge for a data quality project, you build and maintain a knowledge base (KB) that DQS can use to identify incorrect or invalid data. DQS enables you to use both computer-assisted and interactive processes to create, build, and update your knowledge base. Knowledge in a knowledge base is maintained in domains, each of which is specific to a data field. The knowledge base is a repository of knowledge about your data that enables you to understand your data and maintain its integrity.

The following illustration figure 1 displays various components in a knowledge base and a domain in DQS:

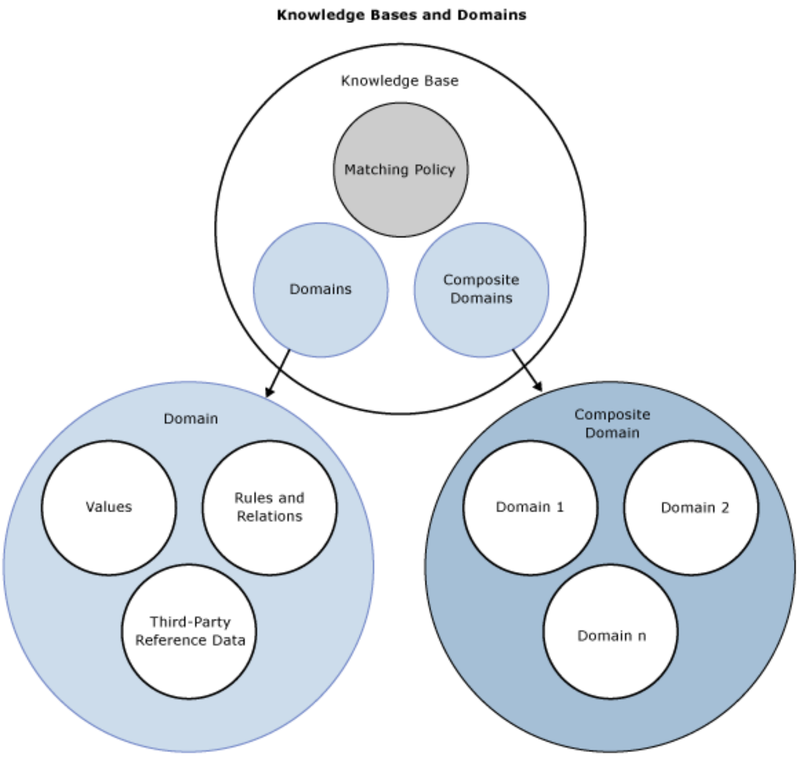


Figure 2. Various components in a knowledge base and a domain in DQS

DQS knowledge bases have the following benefits:

– Building knowledge about data is a detailed process. The DQS process of extracting knowledge about data automatically, from sample data, makes the process much easier

- DQS enables you to see its analysis of the data, and to augment the knowledge in the knowledge base by creating rules and changing data values. You can do so repeatedly to improve the knowledge over time.
- You can leverage pre-existing data quality knowledge by basing a knowledge base on an existing KB, importing domain knowledge from files into the KB, importing knowledge from a project back into a KB, or using the DQS default KB, DQS Data.
- You can ensure the quality of your data by comparing it to the data maintained by a reference data provider.

There is a clear separation between building a knowledge base and applying it in the data correction process, which gives you flexibility in how you build and update the knowledge base.

The data steward uses the Data Quality Client application both to execute and control the computer-assisted steps, and to perform the interactive steps.The advent of online social networks has generated a renaissance in the study of social behaviors and in the understanding of the topology of social interactions. For the first time, it has become possible to analyze networks and social phenomena on a world-wide scale and to design large scale experiments on them. This new evolution in social science has been the center of much attention, but has also attracted a lot of critiques; in particular, a longstanding problem in the study of online social networks is to understand the similarity between them and \real" underlying social networks. This question is particularly challenging because online social networks are often just a realization of a subset of real social networks. For example, Facebook \friends" are a good representation of the personal acquaintances of a user, but probably a poor representation of her working contacts, while LinkedIn is a good representation of work contacts but not a very good representation of personal relationships. Therefore, analyzing social behaviors in any of these networks has the drawback that the results would only be partial. Furthermore, even if certain behavior can be observed in several networks, there are still serious problems because there is no systematic way to combine the behavior of a special user across different social networks and because some social relationships will not appear in any social network. For these reasons, identifying all the accounts belonging to the same individual across different social services is a fundamental step in the study of social science. Interestingly, the problem has also very important practical implications. First, having a deeper understanding of the characteristics of a user across different networks helps to construct a better portrait of her, which can be used to serve personalized content or advertisements. In addition, having information about connections of a user across multiple networks would make it easier to construct tools such as \friend suggestion" or \people you may want to follow". The problem of identifying users across online social networks (also referred to as the social network reconciliation problem) has been studied extensively using machine learning techniques; several heuristics have been proposed to tackle it. However, to the best of our knowledge, it has not yet been studied formally and no rigorous results have been proved for it. One of the contributions of our work is to give a formal definition of the problem, which is a precursor to mathematical analysis. Such a definition requires two key components: A model of the \true" underlying social network, and a model for how each online social network is formed as a subset of this network. Another possible reason for the lack of mathematical analysis is that natural definitions of the problem are demotivating similar to the graph isomorphism problem. In addition, at first sight the social network reconciliation problem seems even harder because we are not looking just for isomorphism but for similar structures, as distinct social networks are not identical. Fortunately, when reconciling social networks, we have two advantages over general graph isomorphism: First, real social networks are not the adversarial designed graphs which are hard instances of graph isomorphism, and second, a small fraction of social network users explicitly link their accounts across multiple networks. The main goal of this paper is to design an algorithm with provable guarantees that is simple, parallelizable and robust to malicious users. Our first contribution is to give a formal model for the graph reconciliation problem that captures the hardness of the problem and the notion of an initial set of trusted links identifying users across different networks. Intuitively, our model postulates the existence of a true underlying graph, then randomly generates realizations of it which are perturbations of the initial graph, and a set of trusted links for some users. Given this model, our next significant contribution is to design a simple, parallelizable algorithm and to prove formally that our algorithm solves the graph reconciliation problem if the underlying graph is generated by well-established network models. It is important to note that our algorithm relies on graph structure and the initial set of links of users across different networks in such a way that in order to circumvent it, an attacker must be able to have a lot of friends in common with the user under attack. Thus it is more resilient to attack than much of the previous work on this topic. Finally, we note that any mathematical model is, by necessity, a simplification of reality, and hence it is important to empirically validate the effectiveness of our approach when the assumptions of our models are not satisfied. Here also remark that for various applications, it may be possible to improve on the performance of our algorithm by adding heuristics based on domain-specific knowledge. For example, we later discuss identifying common Wikipedia articles across languages; in this setting, machine translation of article titles can provide an additional useful signal. However, an important message of this paper is that a simple, efficient and scalable algorithm that does not take any domain-specific information into account can achieve excellent results for mathematically sound reasons.

1.3. A conception of the Semantic Web Expert System and Cloud computing.

A conception of the Semantic Web Expert System which is the logical continuation of the expert system development. The Semantic Web Expert

System emerges as the result of evolution of expert system concept and it means expert system moving toward the Web and using new Semantic Web technologies. The proposed conception of the Semantic Web Expert System promises to have new useful features that distinguish it from other types of expert systems.

Static and dynamic expert systems. The long period of development of expert systems generated different types of expert systems. Despite this fact, it is possible to select a typical structure of a static expert system [15]. Such a structure consists of the following elements:

– Working memory;
– Knowledge base;
– Inference engine;
– Knowledge acquisition component;
– Explanation component;
– Dialogue component.

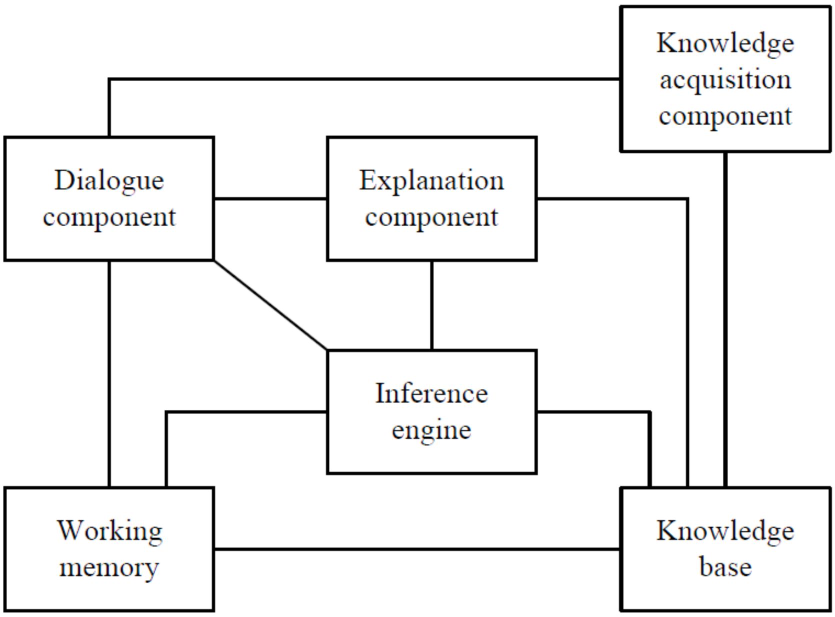


Figure 1.3. The typical structure of a static expert system

It may be sensible to remind the purpose of each part of an expert system:

- Working memory is necessary for data storing which is used for current task solving;
- Knowledge base is necessary to store knowledge, which describes a domain;
- Inference engine is a program that models expert's style of reasoning using knowledge from the knowledge base;
- Knowledge acquisition component automates the process of an expert system filling with data which is executed by an expert;
- Explanation component explains how the system executes the task solution and what knowledge that can facilitate testing and increase trust in the results is used;
- Dialogue component is focused on interaction with users to give the possibility of knowledge input and to show the results of task solution.

CLOUD computing is an emerging paradigm in the IT and data processing communities. Enterprises utilize cloud computing service to outsource data maintenance, which can result in significant financial benefits. Businesses store and access data at remote locations in the "cloud." As the popularity of cloud computing grows, the service providers' face ever increasing challenges. They have to maintain huge quantities of heterogeneous data while providing efficient information retrieval. Thus, the key emphasis for cloud computing solutions is scalability and query efficiency. At the same time semantic web is also an emerging area to augment human reasoning. Resource Description Framework (RDF) which is semantic web technology that can be utilized to build efficient and scalable systems for Cloud Computing. A framework that can be built using Hadoop to store and retrieve large numbers of RDF triples by exploiting the cloud computing paradigm. To determine the Hadoop jobs, and present an algorithm to generate query plan, whose worst case cost is bounded, based on a greedy approach to answer a SPARQL Protocol and RDF Query Language (SPARQL) query. In the project, Hadoop's Map/Reduce framework is used to answer the queries. the results show that it can store large RDF graphs in Hadoop clusters built with cheap commodity class hardware.

Semantic web technologies are being developed to present data in standardized way such that such data can be retrieved and understood by both human and machine. Historically, WebPages are published in plain html files which are not suitable for reasoning. Instead, the machine treats these html files as a bag of keywords. Researchers are developing Semantic web technologies that have been standardized to address such inadequacies. The most prominent standards are Resource Description Framework (RDF) and SPARQL Protocol and RDF Query Language (SPARQL). RDF is the standard for storing and representing data and SPARQL is a query language to retrieve data from an RDF store. Cloud Computing systems can utilize the power of these Semantic web technologies to provide the user with capability to efficiently store and retrieve data for data intensive applications.

The main disadvantage with distributed systems is that they are optimized for relational data. They may not perform well for RDF data, especially because RDF data are sets of triples6 (an ordered tuple of three components called subject, predicate, and object, respectively) which form large directed graphs. In an SPARQL query, any number of triple patterns (TPs) can join on a single variable8 which makes a relational database query plan complex. Performance and scalability will remain a challenging issue due to the fact that these systems are optimized for relational data schemata and transactional database usage. Hadoop is a distributed file system where files can be saved with replication. In addition, it also contains an implementation of the Map/Reduce [6] programming model, a functional programming model which is suitable for the parallel processing of large amounts of data.

Types of clouds. There are three main types of cloud computing:

1. Public Cloud. A public cloud is one in which the infrastructure and other computational resources that it comprises are made available to the general public over the internet. A public cloud is owned by the provider selling cloud

services and is external to the user's organization. However it is submitted that there is greater risk in terms of data security that anyone who subscribes to the cloud has access to it. Thus public cloud shares the characteristics of a multitenant in nature as the data of one company is necessarily stored along with the data of another company on the public cloud.

2. Private Cloud. A private cloud (internal cloud) is one in which the infrastructure and computational resources is operated exclusively for a particular company or organization. Thus, it aims to provide services to limited number of users behind a firewall. The private cloud can be managed either by the organization/company itself or a third party and may be hosted within the organizations data centre or outside it. The private cloud is reminiscent of an internet, access to which is limited to the personnel of a particular company/organization. A private cloud is usually used by a large company and it offers various applications to upgrade or downgrade the resources as required by them. It must be noted that private cloud does not offer the basic advantage of cloud computing because the user still has to incur the up-front capital costs in creating its own private cloud, but these cost is much lesser than the traditional way of owning IT infrastructure.

3. Hybrid Cloud. A hybrid cloud is a composition of two or more clouds (private or public) that remain separate cloud entities but share certain technology which permits interoperability. The hybrid model means that companies can extend their private cloud network to the public cloud service provider. Apart from above three there are three main kind of service models also which aims to offer services. They are:

A). SaaS (Software as A Service). It is the most widely known and widely used form of cloud computing. It provides all the functions of a sophisticated traditional application to many customers and often thousands of users, but through a Web browser, and not a locally-installed application. Little or no code is running on the Users local computer and the applications are usually tailored to fulfill specific functions. SaaS eliminates customer worries about application servers, storage, application development and related, common concerns of IT. Highest-profile examples are Salesforce.com, Google's Gmail and Apps, instant messaging from AOL, Yahoo and Google, and VoIP from Vonage and Skype.

B) PaaS (Platform as a Service). Delivers virtualized servers on which customers can run existing applications or develop new ones without having to worry about maintaining the operating systems, server hardware, load balancing or computing capacity. These vendors provide APIs or development platforms to create and run applications in the cloud – e.g. using the Internet. Managed Service providers with application services provided to IT departments to monitor systems and downstream applications such as virus scanning for e-mail are frequently included in this category. Well known providers would include Microsoft's Azure, Salesforce's Force.com, Google Maps, ADP Payroll processing, and US Postal Service offerings.

C). IaaS (Infrastructure as a Service). Delivers utility computing capability, typically as raw virtual servers, on demand that customers configure and manage. Here Cloud Computing provides grids or clusters or virtualized servers, networks, storage and systems software, usually (but not always) in a multitenant architecture. IaaS is designed to augment or replace the functions of an entire data center. This saves cost (time and expense) of capital equipment deployment but does not reduce cost of configuration, integration or management and these tasks must be performed remotely. Vendors would include Amazon.com (Elastic Compute Cloud [EC2] and Simple Storage), IBM and other traditional IT vendors.

Chapter I summary

To conclude the first chapter of dissertation thesis, would be good to notice, that you can see theoretical view of the social media and social education area, analysis of social networking, advantages and disadvantages. In first chapter was a problems of a knowledge bases in data quality, also was shown benefits of DQS knowledge bases. In third part of first chapter was described cloud computing, semantic web and expert system.

Chapter II.Development of Social Network Analysis, Algorithms of social network and semantic web.

2.1. Development of Social Network Analysis and types of structures of network.

The field of Social Network Analysis today is the result of the convergence of several streams of applied research in sociology, social psychology and anthropology. Many of the concepts of network analysis have been developed independently by various researchers often through empirical studies of various social settings. For example, many social psychologists of the 1940s found a formal description of social groups useful in depicting communication channels in the group when trying to explain processes of group communication. Already in the mid-1950s anthropologists have found network representations useful in generalizing actual field observations, for example when comparing the level of reciprocity in marriage and other social exchanges across different cultures.

Some of the concepts of network analysis have come naturally from social studies. In an influential early study at the Hawthorne works in Chicago, researchers from Harvard looked at the workgroup behavior (e.g. communication, friendships, helping, controversy) at a specific part of the factory, the bank wiring room [Mayo, 1933]. The investigators noticed that workers themselves used specific terms to describe who is in "our group". The researchers tried to understand how such terms arise by reproducing in a visual way the group structure of the organization as it emerged from the individual relationships of the factory workers. In another study of mixed-race city in the Southern US researchers looked at the network of overlapping "cliques" defined by race and age [Warner and Lunt, 1941].3 They also went further than the Hawthorne study in generating hypotheses about the possible connections between cliques. (For example, they noted that lower-class members of a clique are usually only able to connect to higher-class members of another clique through the higher-class members of their own clique.) Despite the various efforts, each of the early studies used a different set of concepts and different methods of representation and analysis of social networks. However, from the 1950s network analysis began to converge around the unique world view that distinguishes network analysis from other approaches to sociological research. (The term "social network" has been introduced by Barnes in 1954.) This convergence was facilitated by the adoption of a graph representation of social networks usually credited to Moreno. What Moreno called a sociogram was a visual representation of social networks as a set of nodes connected by directed links. The nodes represented individuals in Moreno's work, while the edges stood for personal relations. However, similar representations can be used to depict a set of relationships between any kind of social unit such as groups, organizations, nations etc.

While 2D and 3D visual modelling is still an important technique of network analysis, the sociogram is honored mostly for opening the way to a formal treatment of network analysis based on graph theory. The following decades have seen a tremendous increase in the capabilities of network analysis mostly through new applications. SNA gains its relevance from applications and these settings in turn provide the theories to be tested and greatly influence the development of the methods and the interpretation of the outcomes. For example, one of the relatively new areas of network analysis is the analysis of networks in entrepreneurship, an active area of research that builds and contributes to organization and management science. The vocabulary, models and methods of network analysis also expand continuously through applications that require to handle ever more complex data sets. An example of this process are the advances in dealing with longitudinal data. New probabilistic models are capable of modelling the evolution of social networks and answering questions regarding the dynamics of communities. Formalizing an increasing set of concepts in terms of networks also contributes to both developing and testing theories in more theoretical branches of sociology.

The global structure of networks. As discussed above, a (social) network can be represented as a graph $G = (V;E)$ where V denotes the finite set of vertices and E

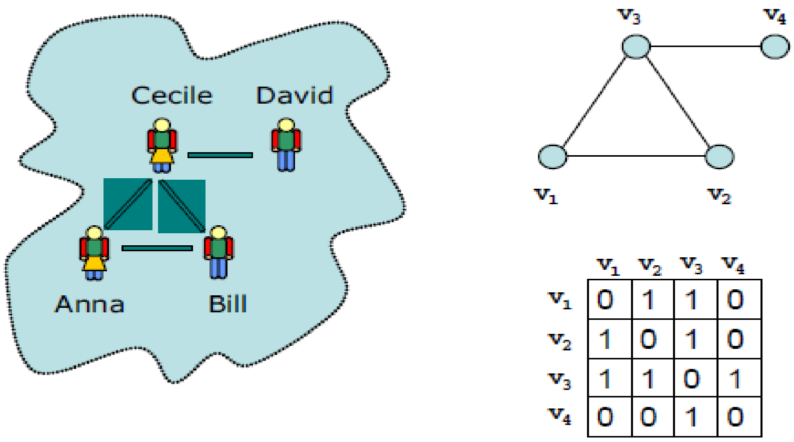
$$n = |V|, m_{i,j} = \left\{ \begin{array}{c|c} 1 & (v_i, v_j) \in E \\ 0 & otherwise \end{array} \right\}$$

denotes a finite set of edges such that *E µ V £V* . Recall that each graph can be associated with its characteristic matrix *M* := (*mi;j*)*n*¤*n* where

Some network analysis methods are easier to understand when we conceptualize graphs as matrices (see Figure 2.1). Note that the matrix is symmetrical in case the edges are undirected. We will talk of a valued graph when we are also given a real valued weight function *w*(*e*) defined on the set of edges, i.e. *w*(*e*) := *E £* R. In case of a valued graph, the matrix is naturally defined as

$$m_{i,j} = \left\{ \begin{array}{c|c} w(e) & (v_i, v_j) \in E \\ 0 & otherwise \end{array} \right\}$$

Loops are not excluded in the above definition, although they rarely in occur in practical social network data sets. (In other words, the main diagonal of the matrix is usually empty.) Typically, we also assume that the network is connected, i.e. there is a single *(weak) component* in the graph.6 Otherwise we choose only one of the components for analysis.



	v1	v2	v3	v4
v1	0	1	1	0
v2	1	0	1	0
v3	1	1	0	1
v4	0	0	1	0

Figure 2.1. Most network analysis methods work on an abstract, graph based representation of real world networks.

At this point the question of what else can we say about the structure of social networks arises naturally. Are there any commonalities to real world (social) networks that we could impose on our graph models or any kind of graph is equally likely to occur in practice? These questions are relevant because having a general model of social networks would allow us to answer questions in general, i.e. in a way that the answer should hold for all networks exhibiting the common characteristics. In some cases we could verify theories solely on an abstract model instead of having to collect network data. Further, understanding the global structure of networks can lead us to discover commonly occurring patterns of relationships or typical network positions worthy of formalizing in further detail.

Milgram calculated the average of the length of the chains and concluded that the experiment showed that on average Americans are no more than six steps apart from each other. While this is also the source of the expression six degrees of separation the actual number is rather dubious: not only was Milgram's sample too small, but even only 20% of the those letters have made it to their destination. Thus the number could be actually larger: those letters that did not make it would probably have resulted in longer paths. But the number could be also smaller as it is not guaranteed that the letters have travelled the shortest possible path from their source to the target. Still, Milgram's experiment had a tremendous impact on social network research and sociology as a whole as it showed that the number is orders of magnitude smaller than the size of the network.

A practical impact of Milgram's finding is that we can exclude certain kind of structures as possible models for social networks. The two dimensional lattice model shown in Figure 2.2, for example, does not have the small world property: for a network of size n the characteristic path length is $2=3 \times \sqrt{n}$, which is still too large a number to fit the empirical finding.

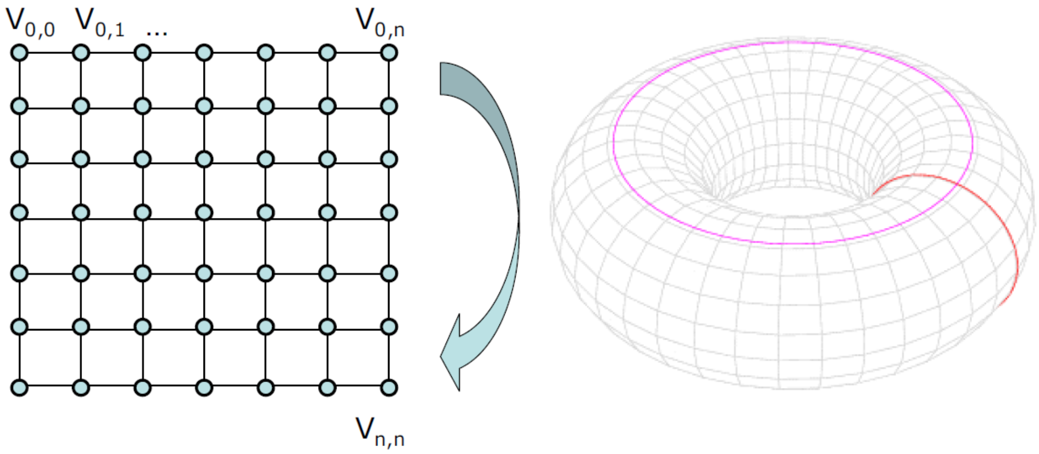


Figure 2.2. The 2D lattice model of networks (left). By connecting the nodes on the oppositeborders of the lattice we get a toroidal lattice (right).

The macro-structure of social networks. Based on the findings about the global characteristics of social networks we now have a good impression about what they might look like. In particular, the image that emerges is one of dense clusters or social groups sparsely connected to each other by a few ties as shown in Figure 2.3. (These weak ties have a special significance as we will see in the following Section.) For example, this is the image that appears if we investigate the co-authorship networks of a scientific community. Bounded

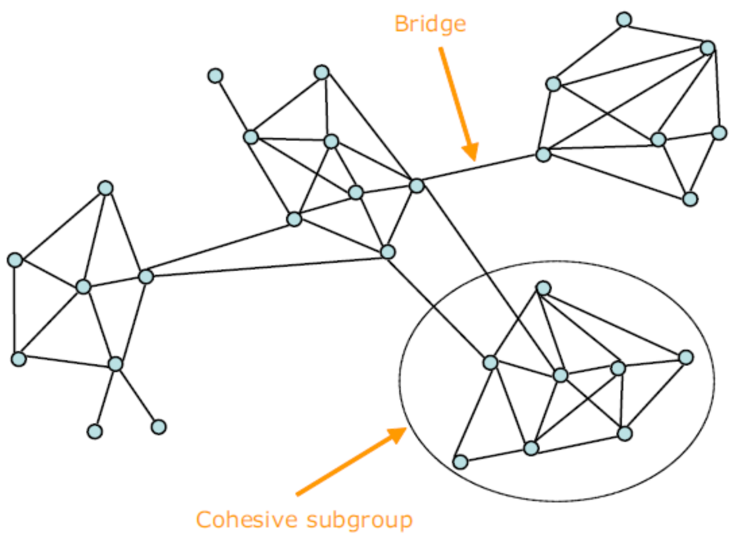


Figure 2.3. Most real world networks show a structure where densely connected subgroups are linked together by relatively few bridges.

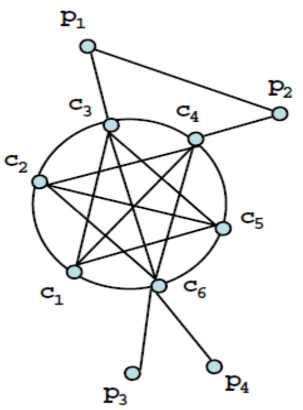
by limitations of space and resources, scientists mostly co-operate with colleagues from the sameinstitute. Occasional exchanges and projects with researchers from abroad, however, create the kind of shortcut ties that Watts explicitly incorporated within his model. These short cuts make it possible for scientists to reach each other in a relatively short number of steps.

Particularly important are the hubs, the kind of individual such as Erd½os was in the field of mathematics: he was one of the most prolific authors in the history of mathematics and co-authored with an unusual number of individuals.

Clustering a graph into subgroups allows us to visualize the connectivity at a group level. In some cases we already have an idea of what this macro-structure might look like. A typical pattern that often emerges in social studies is that of a Core-Periphery (C/P) structure. A C/P structure is one where nodes can be divided in two distinct subgroups: nodes in the core are densely connected with each other and the nodes on the periphery, while peripheral nodes are not connected with each other, only nodes in the core (see Figure 2.4.). $\begin{pmatrix} 1 & \cdot \\ \cdot & 0 \end{pmatrix}$

The matrix form of a core periphery structure is a

Algorithms for identifying C/P structures and other block models (structural patterns) work by dividing the set of nodes in a way that the error the between



c_1	c_2	c_3	c_4	c_5	c_6	p_1	p_2	p_3	p_4
1	1	1	1	1	1	0	0	0	0
1	1	1	1	1	1	0	0	0	0
1	1	1	1	1	1	1	0	0	0
1	1	1	1	1	1	0	1	0	0
1	1	1	1	1	1	0	0	0	0
1	1	1	1	1	1	0	0	1	1
0	0	1	0	0	0	0	1	0	0
0	0	0	1	0	0	1	0	0	0
0	0	0	0	0	1	0	0	0	0
0	0	0	0	0	1	0	0	0	0

Figure 2.4. A Core-Periphery structure that would be perfect without the edge between nodes $p_1$and p_2.

the actual image and the “perfect” image is minimal. The result of the optimization is a classification of the nodes as core or periphery and a measure of the error of the solution.

Encounter situations in our work where it is some additional information that allows us to group our nodes into categories. For example, in the case of a scientific community we might have data on the interests or affiliations of researchers.

2.2. Social networking system algorithm

The problem of identifying Internet users was introduced to identify users across different chat groups or web sessions in J. Novak, P. Raghavan, and A. Tomkins. Anti-aliasing on the web. In WWW, and J. R. Rao and P. Rohatgi. Can pseudonymity really guarantee privacy? In USENIX. Both papers are based on similar intuition, using writing style (stylography features) and a few semantic features to identify users. The social network reconciliation problem was introduced more recently by Zafarani and Liu. The main intuition behind their paper is that users tend to use similar usernames across multiple social networks, and even when different, search engines and the corresponding names. To improve on these first naive approaches, several machine learning models were developed, all of which collect several features of the users (name, location, image, connections topology), based on which they try to identify users across networks. These techniques may be very fragile with respect to malicious users, as it is not hard to create a fake profile with similar characteristics. Furthermore, they get lower precision experimentally than our algorithm achieves. However, we note that these techniques can be combined with ours, both to validate / increase the number of initial trusted links, and to further improve the performance of our algorithm. A different approach was studied, where the authors infer missing attributes of a user in an online social network from the attribute information provided by other users in the network. To achieve their results, they retrieve communities, identify the main attribute of a community and then spread this attribute to all the user in the community. Though it is interesting, this approach suffers from the sameΔΔ limitations of the learning techniques discussed above. Recently, Henderson et al. studied which are the most important features to identify a node in a social network, focusing only on graph structure information. They analyzed several features of each ego-network, and also added the notion of recursive features on nodes at distance larger than from a specific node. It is interesting to notice that their recursive features are more resilient to attack by malicious users, although they can be easily circumvented by the attacker typically assumed in the social network security literature, who can create arbitrarily many nodes. The problem of reconciling social networks is closely connected to the problem of de-anonymizing social networks. Back strom et al. introduced the problem of deanonymizing social networks. In their paper, they present main techniques: An active attack (nodes are added to the network before the network is anonymized), and a second passive one. Our setting is similar to that described in their passive attack. In this setting the authors are able to design a heuristic with good experimental results; though their technique is very interesting, it is somewhat elaborate and does not have a provable guarantee. In the context of de-anonymizing social networks, the work of Narayanan and Shmatikov is closely related. Their algorithm is similar in spirit to ours; they look at the number of common neighbors and other statistics, and then they keep all the links above a specific threshold. There are two main differences between our work and theirs. First, we formulate the problem and the algorithm mathematically and we are able to prove theoretical guarantees for our algorithm. Second, to improve the precision of their algorithm the authors construct a scoring function that is expansive to compute. In fact the complexity of their algorithm is $O((E1 + E2)\Delta 1\Delta 2)$, where E1 and E2 are the number of edges in the two graphs and $\Delta 1$ and $\Delta 2$ are the maximum degree in the 2 graphs. Thus their algorithm would be too slow to run on Twitter and Facebook, for example; Twitter has more than 200M users, several of whom have degree more than 20M and Facebook more than 1B users with several users of degree 5K. Instead, in our work we are able to show that a very simple technique based on degree bucketing combined with the number of common neighbors succes to guarantee strong theoretical guarantees and good experimental results. In this way we designed an algorithm with sequential complexity O((E1+E2)min(Δ1;Δ2) log(max(Δ1;Δ2))) that can be run in O(log(max(Δ1;Δ2))) MapReduce rounds. In this context, our paper can be seen as the first really scalable algorithm for network de-anonymization with theoretical guarantees. Further, we also obtain considerably higher precision experimentally, though a perfect comparison across different datasets is not possible. The different contexts also are important: In de-anonymization, the precision of 72% they report corresponds to a significant violation of user privacy. In contrast, we focus on the benefits to users of linking accounts; in a user-facing application, suggesting and account with a 28% chance of error is unlikely to be acceptable.

Model. Recall that a formal definition of the user identification problem requires first a model for the \true" underlying social network G(V,E) that captures relationships between people. However, we cannot directly observe this network; instead, we consider two imperfect realizations or copies G1(V,E1) and G2(V,E2) with E1;E2 C E. Second, we need a model for how edges of E are selected for the two copies E1 and E2. This model must capture the fact that users do not necessarily replicate their entire personal network s on any social networking service, but only a subset. Any such mathematical models are necessarily imperfect descriptions of reality, and as models become more `realistic', they become more mathematically intractable. In this paper, we consider certain well-studied models, and provide complete proofs. It is possible to generalize our mathematical techniques to some variants of these models; for instance, with small probability, the two copies could have new \noise" edges not present in the original network G(V,E), or vertices could be deleted in the copies. We do not fully analyze these as the generalizations require tedious calculations without adding new insights. For the underlying social network, our main focus is on the preferential attachment model ,which is historically the most cited model for social networks. Though the model does not capture some features of real social networks, the key properties we use for our analysis are those common to online social networks such as a skewed degree distribution, and the fact that nodes have distinct neighbors including some long-range / random connections not shared with those immediately around them. In the experimental section we will consider also different models and also real social networks as our underline real networks. For the two imperfect copies of the underlying network we assume that G1 (respectively G2) is created by selecting each edge e Є E of the original graph G(V;E) independently with a fixed probability s1 (resp. s2) (See Figure 1.) In the real world, edges/relationships are not selected truly independently, but this serves as a reasonable approximation for observed networks. In fact, a similar model has been previously considered, which also produced experimental evidence from an email network to support the independent random selection of edges. Another plausible mechanism for edge creation in social network is the cascade model, in which nodes are more likely to join a new network if more of their friends have joined it. Experimentally, we show that our algorithm performs even better in the cascade model than in the independent edge deletion model. These two models are theoretically interesting and practically interesting. Nevertheless, in some cases the analyzed social networks may differ in their scopes and so the group of friends that a user has in a social network can greatly differ from the group of friends that same user has in the other network. To capture this scenario in the experimental section, we also consider the Affiliation Network model as the underlying social network. For each of G1;G2, and for each community, we keep or delete all the edges inside the community with constant probability. This highly correlated edge deletion process captures the fact that a user's personal friends might be connected to her on one network, while her work colleagues are connected on the second network. We defer the detailed description of this experiment to Section 5. Recall that the user identification problem, given only the graph information, is intractable in general graphs. Even the special case where s1 = s2 = 1 (that is, no edges have been deleted) is equivalent to the well-studied Graph Isomorphism problem, for which no polynomial-time algorithm is known. Fortunately, in reality, there are additional sources of information which allow one to identify a subset of nodes across the two networks: For example, people can use the same email address to sign up on multiple websites. Users often explicitly connect their network accounts, for instance by posting a link to their Facebook profile page on Google+ or Twitter and vice versa. To model this, we assume that there is a set of users/nodes explicitly linked across the two networks G1;G2. More formally, there is a linking probability l (typically, l is a small constant) and each node in V is linked across the networks independently with probability l. (In real networks, nodes may be linked with differing probabilities, but high-degree nodes / celebrities may be more likely to connect their accounts and engage in cross-network promotions; this would be more likely to help our algorithm, since low-degree nodes are less valuable as seeds because they help identify only a small number of neighbors. In the experiments of, the authors explicitly consider high-degree nodes as seeds in the real-world experiments.) A natural algorithm to solve the user identification problem with a set of linked nodes, and discuss some of its properties. This algorithm performs well on several well established network models. The algorithm also works very well in practice, by examining its performance on real-world networks.

The Algorithm. To solve the user identification problem, we design a local distributed algorithm that uses only structural information about the graphs to expand the initial set of links into a mapping/identification of a large fraction of the nodes in the two networks.

Before describing the algorithm, we introduce a useful definition.

Definition 1. A pair of nodes (u1; u2) with u1 Є G1; u2 Є G2 is said to be a similarity witness for a pair (v1; v2) with v1 Є G1; v2 Є G2 if u1 2 N1(v1); u2 Є N2(v2) and u1 has been linked to / identified with u2.

Here, N1(v1) denotes the neighborhood of v1 in G1, and similarly N2(v2) denotes the neighborhood of v2 in G2. Roughly speaking, in each phase of the algorithm, every pair of nodes (one from each network) computes and similarity score that is equal to the number of similarity witnesses they have. We then create a link between two nodes v1 and v2 if v2 is the node in G2 with maximum similarity score to v1 and vice versa. We then use the newly generated set of links as input to the next phase of the algorithm. A possible risk of this algorithm is that in early phases, when few nodes in the network have been linked, low-degree nodes could be mis-matched. To avoid this (improving precision), in the phase, we only allow nodes of degree roughly D=2i and above to be matched, where D is a parameter related to the largest node degree. Thus, in the first phase, we match only the nodes of very high degree, and in subsequent phases, we gradually decrease the degree threshold required for matching. In the experimental section we will show in fact that this simple step is very effective, reducing the error rate by more than 33%. We summarize the algorithm, that called User-Matching, as follows:

Input:

$G_1(V,E_1)$;G2(V,E_2);L a set of initial identification links across the networks, D the maximum degree in the graph a minimum matching score T and a specified number of iteration k.

Output:

A larger set of identification links across the networks.

Algorithm:

For i = 1,… , k

For j = log D,…, 1

For all the pairs (u, v) with u Є G_1 and v Є G_2 and such that d_{G1} (u) ≥ 2^j and d_{G2} (v) ≥ 2^j

Assign to (u; v) a score equal to the number of similarity witnesses between u and v If (u; v) is the pair with highest score in which either u or v appear and the score is above T add (u; v) to L.

Output L

Where $d_{G_i}(u)$ is the degree of node u in G_i. Note that the internal for loop can be implemented efficiently with 4 consecutive rounds of MapReduce, so the total running time would consist of O(k logD) MapReductions. In the experiments, we note that even for a small constant k (1 or 2), the algorithm returns very interesting results. The optimal choice of threshold T depends on the desired precision/recall tradeo; higher choices of T improve precision, but in our experiments, we note that T = 2 or 3 is sufficient for very high precision.

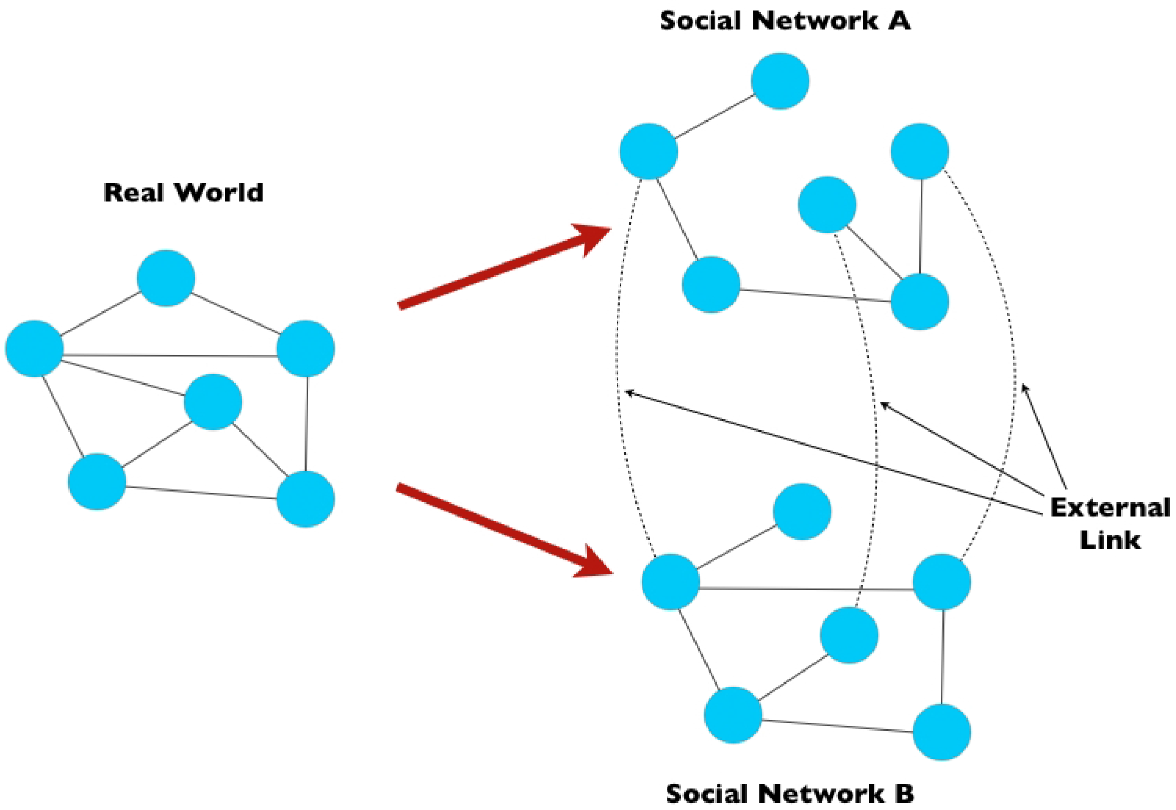


Figure 2.6. From the real underlying social network, the model generates two random realizations of it, A and B, and some identification links for a subset of the users across the two realizations.

2.3. Semantic Matchmaking Algorithm

The ability to dynamically discover and invoke a Web Service is a critical aspect of Service Oriented Architectures. An important component of the discovery process is the matchmaking algorithm itself. In order to overcome the limitations of a syntax-based search, matchmaking algorithms based on semantic techniques have been proposed. Most of them are based on an algorithm originally proposed by M. Paolucci, et al. [21].

Loose Coupling is an important principle underlying Service Oriented Architectures. One aspect of loose coupling is the ability to invoke a service provider with little (or no) prior knowledge about it. The publish-find-bind architecture is intended to facilitate this process. Service providers create WSDL [9] descriptions and publish them to UDDI [8] registries. Clients search the registry to locate providers of the desired service. Today, in most cases, the WSDL is compiled into client-stubs and the service is invoked. This approach, however, has several limitations. The WSDL is a specification of the messaging syntax between the client and the provider. It is necessary for a human to interpret the WSDL and then invoke the client stub with the correct parameters. The search capabilities of UDDI are limited to a syntax based search. A client can search the registry for a string in the service description or it can search the service classification hierarchy (like NAICS [3]) in the TModel. Neither of these techniques are sufficient, for a client, to be able to autonomously choose a service provider and invoke it without human intervention. In order to overcome these limitations, techniques for semantic description and matchmaking of services have been proposed in recent literature. These techniques use semantic concepts from Ontologies to describe the Inputs, Outputs, Pre-conditions and Effects (IOPE) of a service. The discovery process involves the matchmaking of the semantic descriptions offered by the client and the provider. In this paper we analyze the semantic matchmaking algorithm proposed by Paolucci, et al. [21]. We have considerable interest in this algorithm because it has been cited extensively in recent literature and several subsequent proposals ([13], [22], [16], [14]) are based on it. The outline of the paper is as follows: First, we present the algorithm by Paolucci [21]. We then present counterexamples where this algorithm does not generate correct outcomes. We describe our own matchmaking algorithm which overcomes these correctness issues. Finally, we analyze the complexity of the two algorithms and present some experimental results in order to compare their performance.

The algorithm takes a OWL-S Query from the client as input and iterates over every OWL-S Advertisement in its repository in order to determine a match. An Advertisement and a Query match if their Outputs and Inputs, both, match. The algorithm returns a set of matching advertisements sorted according to the degree of match.

The algorithm takes a OWL-S Query from the client as input

Iterates over every OWL-S Advertisement in its repository in order to determine a match

An Advertisement and a Query match if their Outputs and Inputs match

The algorithm returns a set of matching advertisements sorted according to the degree of match

Algorithm.

Let Query out and Advt out represent the list of output concepts of the Query and Advertisement respectively

Matching the outputs requires the matching between two concept-lists, Query

$$\forall c \in Query_{out}, \exists d \in Advt_{out}, \\ \text{s.t. } match(c, d) \neq Fail$$

out and Advt out

Let Query in and Advt in represent the list of input Concepts of the Query and Advertisement respectively.

Matching the inputs requires:

$$\forall c \in Advt_{in}, \exists d \in Query_{in}, \\ \text{s.t. } match(c, d) \neq Fail.$$

Note that the order of Query and Advt has been transposed between the two expressions above. Suppose out Q ∈Query out and out A ∈Advt out are two concepts. In case of output matching, the match(out Q, out A) function accepts out Q and out A as inputs and returns the degree of match between them. Four degrees of match are defined between a them:

Exact: If out A is an equivalent concept to out Q or out A is a superclass of out Q. In case of a superclass relationship, it is assumed that the service provider has agreed to support every possible subclass of out A.

Plugin: If out A Subsumes out Q. The relation between out A and out R is weaker as compared to the previous case since subsumption is indirectly inferred by the reasoner. It is assumed that the provider has agreed to support some sub-concepts of out A. We hence infer that out A can be plugged in place of the required out R.

Subsume: If out Q Subsumes out A. The set of individuals defined by the concept, out A, is a subset of the set of individuals defined by the concept out R.

Fail: If none of the above conditions are satisfied. These four degrees as ranked as: Exact >Plugin >Subsumes >Fail. Here, x > y indicates that x is ranked higher (is a more desirable match) than y.

Degrees of match. Four degrees of match are defined between a them:

Exact:

If outA is an equivalent concept to outQ or outA is a superclass of outQ

In case of a superclass relationship,

it is assumed that the service provider has agreed to support every possible subclass of outA

Plugin:

If outA Subsumes outQ

The relation between outA and outR is weaker as compared to the previous case since subsumption is indirectly inferred by the reasoner

It is assumed that the provider has agreed to support some sub-concepts of outA

Hence infer that outA can be plugged in place of the required outR

Chapter II summary

The main results of this chapter are:

- Researched structure development of social networking;
- Analyzed the types of network structure with figure where was shown;
- Researched the web semantic and social networking algorithms;
- Described main two types, the global structure of networks and macro structure. There are some sub types of them also was shown.

ChapterIII.Software description. Solving the semantic web and social networking problem with genetic algorithm.

3.1. Functions of AP as a social network site

Automatic composition of web services supports the solving of complex user request. The set of possible solutions can be represented by a graph, modeling the composition. Usually, this kind of approach is highly simplified by considering only sequences of services. Proposes an algorithm for automatic semantic web services composition, which generates a graph taking into account any composition structure. The request resolution process identifies possible composition structures and selects relevant services based on their semantic description.

The study of social networks has received significant interest from researchers in various domains such as psychology, philosophy, education, and lately computer science – particularly in the field of artificial intelligence. This section defines what we mean by social networks, the way in which these networks form and evolve in our daily lives, and their relations to other types of group.

There are different kinds of SNSs developed for many different kinds of purposes. The main common characteristics of these sites are to make their users be able to create profiles, to reveal the list of the people with whom the user is in contact and to access the contact lists of the others. Similarly, the user profile pages on AP have been created in a detailed way. The contents added by the AP members appear on the profile pages specially designed for the user. The contents created by the users can be commented by other users. The users can access the detailed information about a user by visiting the profile page of his/her page who creates contents of news or projects. The visitor, who wants to communicate with the user creating the content, through his/her profile page, can communicate directly with this user by using the e-mail system developed on the site. Besides, through the "report" function, the users can deliver the contents which they consider as harmful, directly to the site administration. AP members can create many different types of UCC such as picture, video and writings. AP has a very flexible structure in terms of user-created contents. The members can produce news about architecture and interior designing, announcements, trends and projects. These contents can be added to the site in script, picture and video formats. All the contents created can be commented by the users. The freedom of the communication of the members has been increased through this feature. Private profile pages have been generated for every single member. The members can create detailed data on their profile pages about contact information, company information, interest fields, information of profession and educational information. AP has detailed search functions which have been designed especially for the members to provide quick access to the information. Through this feature, the members can access the information they search about architecture or interior design faster than every other page. AP currently does not have the functions to create online connections and friend lists. However, the profile pages generated for the users have been designed very flexible and detailed. The functions on AP can differ in terms of individual members and institutional members.

3.2. Architectural platform: Application Platform

The network AP is an information and sharing network which was developed on a frame of purposes that aim to follow architectural developments, to create and manage up to date information databases about too fast changing architectural trends, to connect the visitors to the architecture offices, through its wide architectural database to be the first access point of the architectural information in the sector. The site AP aims to be a meeting point for the producers, architects, interior designers, academicians, students and the other visitors interested in architecture.

The producers and architects can exhibit their products, the students and academicians can follow the architectural developments and use an infinite literature, and the visitors interested in architecture can easily access to all of the architecture. The main page of AP is shown in Figure 3.1.

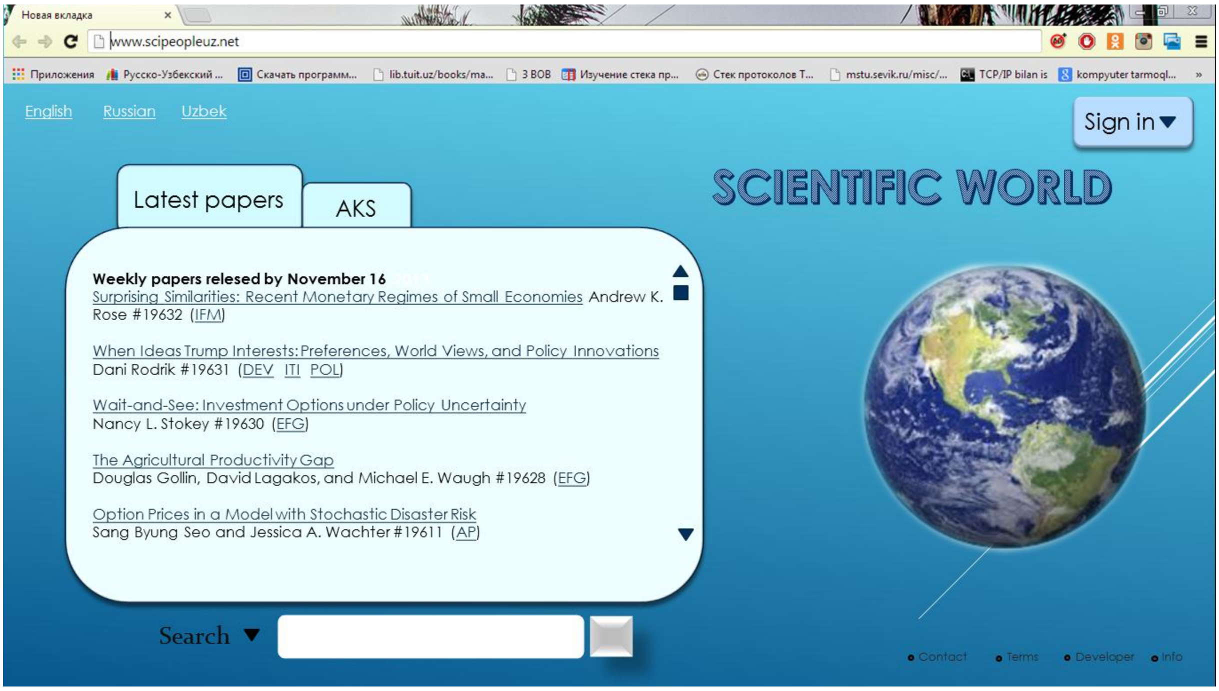


Figure 3.1. Main page

There are different kinds of services that AP provides for different user types. The possibilities have been offered for professionals to exchange their ideas on project solutions, to know new talents through the bulletin board and the display window, to follow the actual developments by following the national and international educationcompetition, fair, seminar, conference and workshop announcements. Moreover, the professional members, by adding contents to news, announcements and trend sections, find the possibility to present themselves to other colleagues and to the ones interested in decoration who visit the site over another site which includes different brands and products and which has been prepared by the experts in their fields.

For the students, the options have been offered to introduce themselves to the leading product providing companies of the sector, to the brands, to the architecture and interior designing offices by publishing their projects and to send internship and job applications to those companies and offices. The offered options for the academics are to exchange ideas with the colleagues who are parts of other national or international educational institutions in the sector and to follow the actual developments by following the national and international education, competition, fair, seminar, conference and workshop announcements. For all the other members, such as the ones interested in decoration or the ones wanting to decorate their houses or the ones who search furniture by walking the streets, such possibilities are offered to decide what to buy before shopping, to get information on the brand of the production which is going to be bought and to save time by reaching the product they want by finding the address on the map of the related providing company.

Typically, institutions use a range of various educational approaches in the classroom, tutorial, lab and lecture hall. Activities can take place face to face, but may also be mediated by social networking technologies include peer assessment, discussions, and collaborative work. Course designers have been quick to spot such opportunities by way of chat rooms, discussion forums and collaborative work support tools which may be used in this way. The efficiency and effectiveness of such approaches are necessarily the subject of evaluation, analysis and debate. The study of social networks within a learning domain encompasses the processes of social learning that occurs when a self-selecting group of people who have a common interest in a subject collaborate to share ideas or find solutions. Observations of the processes and behaviors of self-selecting groups can be used to engineer interactions in groups orchestrated for specific educational purposes. Social networking applications which incorporate Web 2.0 technologies demonstrate affordances, which could be available to utilize within the classroom. These operate with paradigms which are different to those observed within conventional e-learning tools. However utilizing social networking tools with large student groups might present problems. An advantage of increased awareness or appreciation of the complexity of typical observed behaviors in a social learning environment may enhance the academic's ability to manage the tools. A recent study of the potential for semantic modeling of learners explores using Semantic Web-based social networks to facilitate the automatic and dynamic creation of students' networks within large online communities. Enriching the semantics of network and membership descriptions can provide valuable information. This can be used to assist in tuning group allocations, enabling the network to be used for specific educational objectives.

Computer science is a discipline with its own distinct culture, and this affects the nature of social interaction online. As diversity isdiscussed elsewhere in this paper, the shared characteristics of computer science students with their study-mates and the impact of their shared tasks bear consideration. Prentice et al. distinguish between (real life) groups based around a common bond, and those based on a common identity. A common bond group is based primarily on connections that exist among peers – for example in a friendship group. Attachment to a common identity group is based on an individual's identification with a group's purpose. Subsequently established that the distinctions and processes outlined by Prentice et al. also apply to online groups. In universities, entrance criteria are applied to an already self-selected population who have been educated to a similar level, in similar subjects and express similar interests. Thus, students' educational experiences at admission are broadly analogous. The vast majority of undergraduates have grown up in the same globalized technological milieu, and are characterized as a generation of 'millennials' or 'digital natives', for whom computers and the internet are mundane, rather than exceptional. Computer science undergraduates are, by virtue of their specialization, highly computer literate and when forming study mates they can be expected to be initially based upon a common identity. Prentice et al. demonstrates that common identity groups show higher levels of conformity and adherence to group norms than common-bond groups. This suggests that there will be limited diversity in common identity groupssuch as computer science. Subsequentlyestablished that the distinctions and processes outlined byPrentice et al. also apply to online groups. This suggests that one might expect therefore that computer science's owncultural 'norms' would be magnified in curriculum-specific social networks. To approach such generalizations in the case of computer scientists, this research must ensure close attention to the demographics of the student participants involved. The situated relationships between individuals, their social network and institution must be respected. Computer science is a discipline with its own distinct culture, and this affects the nature of social interaction online. As diversity is discussed elsewhere in this paper, the shared characteristics of computer science students with their cohorts and the impact of their shared tasks bear consideration.

3.3. Software structure

Finally, in this part will show structure of social network site. It given in figure 3.2. Given scheme describes main meaning of proposed idea, there was shown hierarchical scheme, where all the steps going step by step. Beginning with main page, here was shown registration steps, choosing disciplines, topics, subject and research area directions, main types of search with sub menus, after

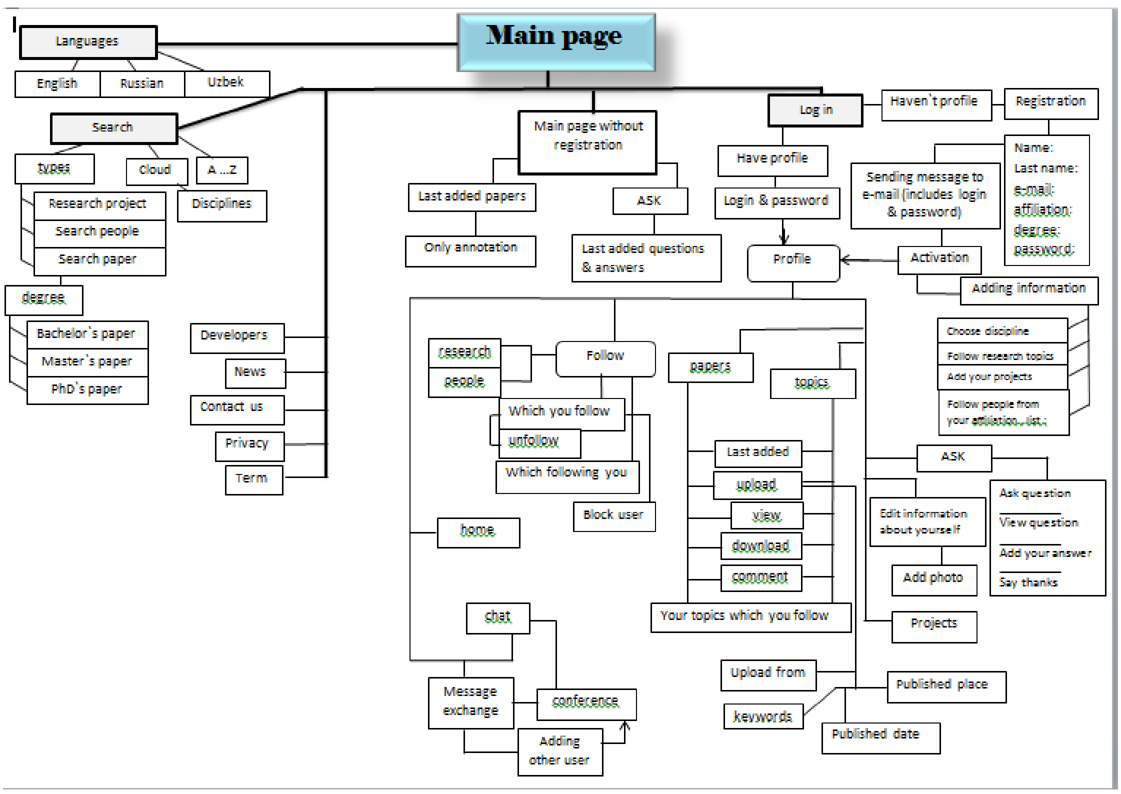


Figure 3.2. Proposed scheme of network

registration profile view main menus of user`s profile such as ask, where anyone can ask some question from everyone, topics, papers.

In the figure 3.3 was shown registration to the site, and here you can see

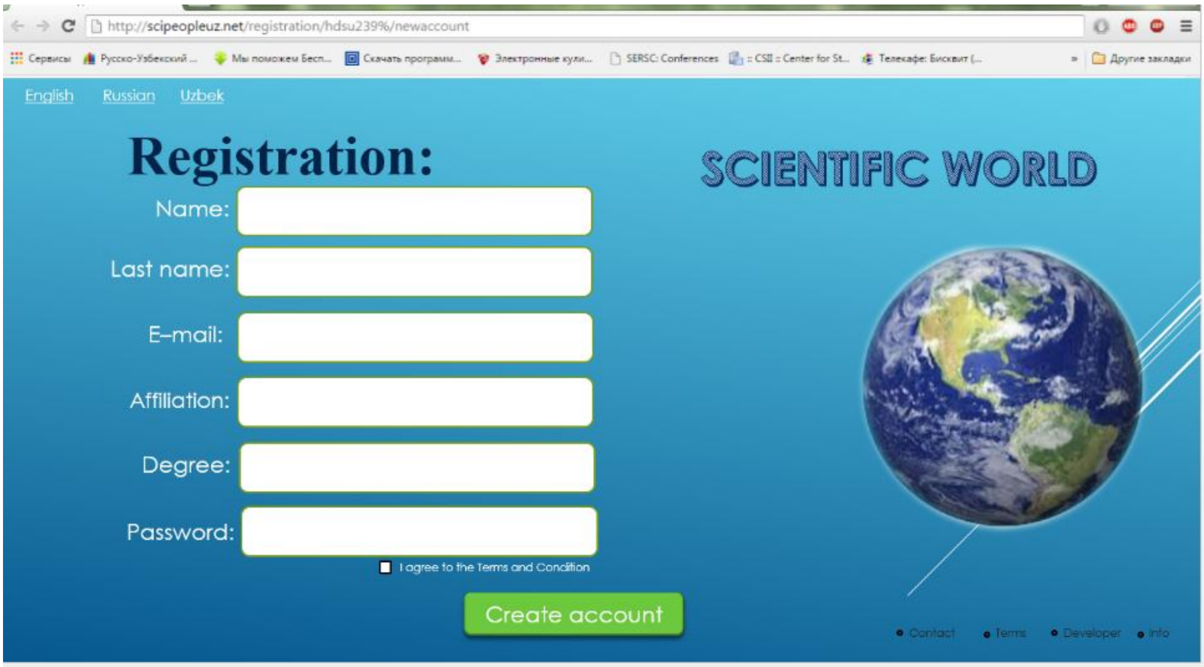


Figure 3.3. Registration to the site

step of it. After registration to user will send a message with activation after activation user can be able to choose need disciplines. It shown in figures 3.5 and 3.6.

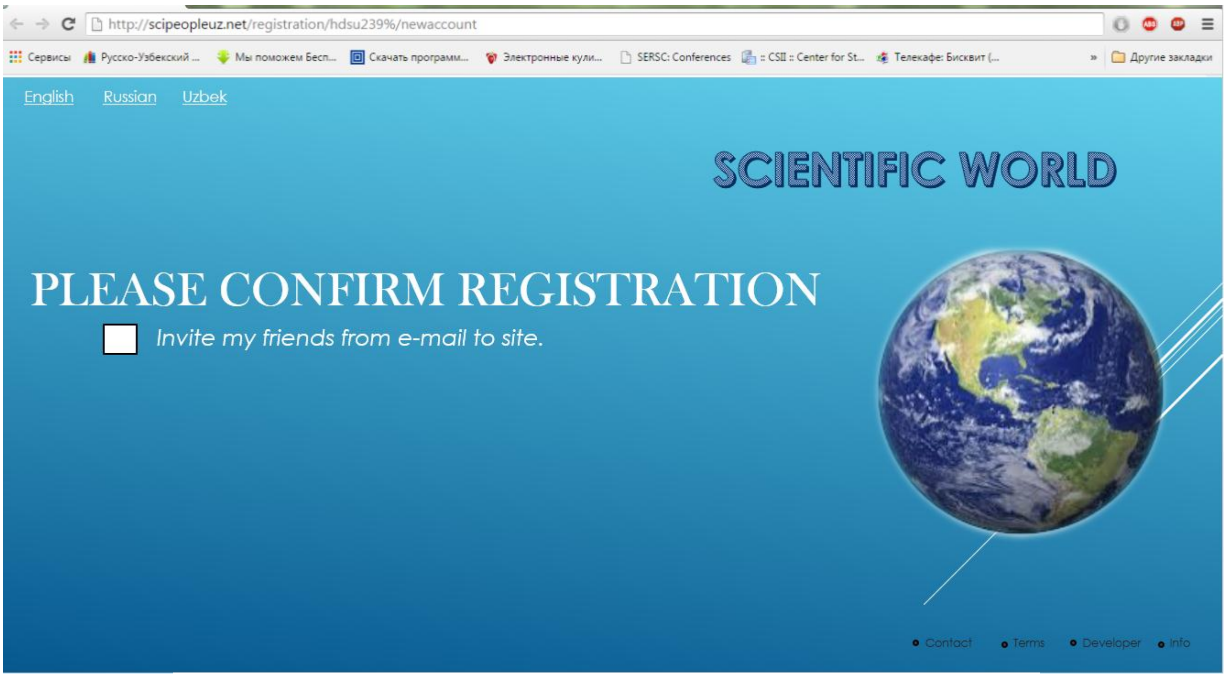


Figure 3.4. Activation window

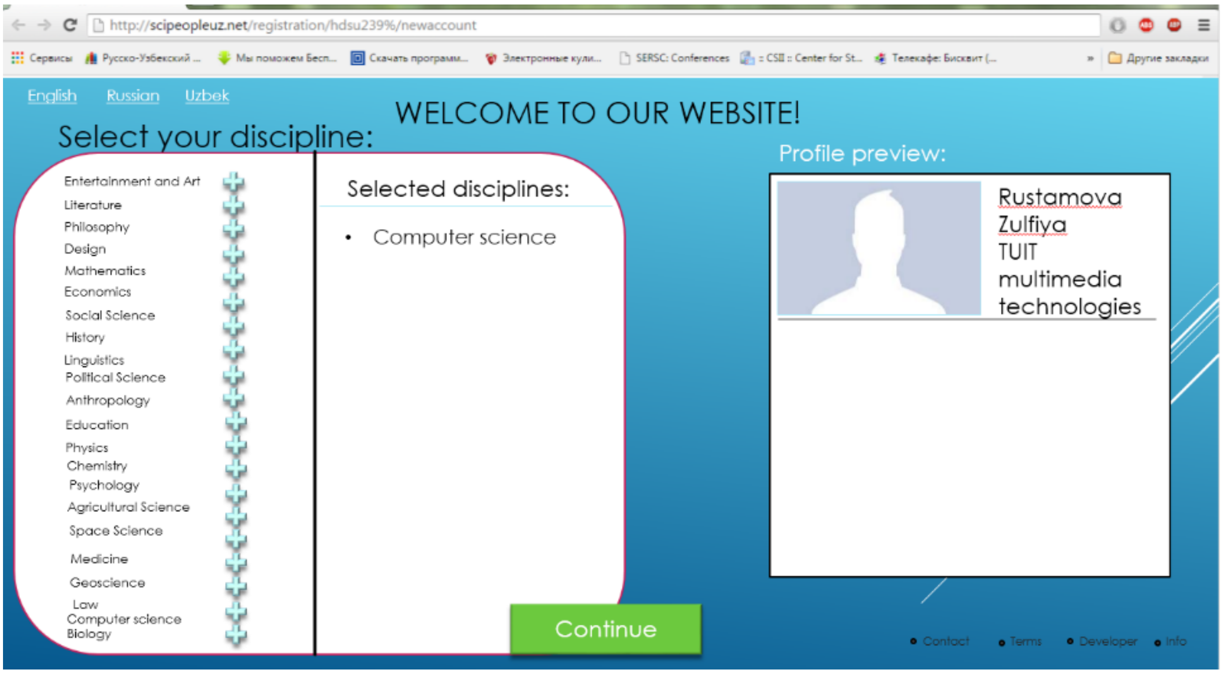


Figure 3.5. Choosing disciplines

After selecting needed discipline, user may select those discipline`s types and select every needed topics (figure 3.6.).

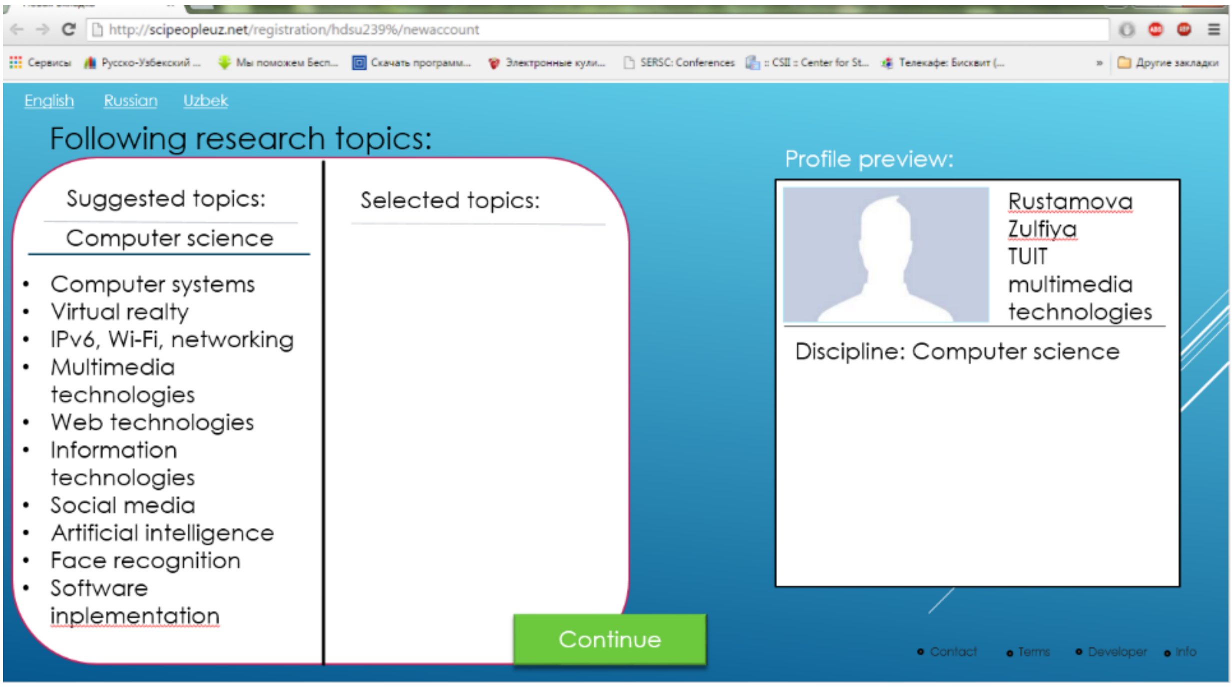


Figure 3.6. Selecting topics

Chapter III summary

Final chapter shows all about software implementation. The study of social networks has received significant interest from researchers in various domains such as psychology, philosophy, education, and lately computer science – particularly in the field of artificial intelligence. This section defines what we mean by social networks, the way in which these networks form and evolve in our daily lives, and their relations to other types of group. As there are many types of social network site, was chose educational site and in this chapter you may see all it with given figures of software platform.

Conclusion

Researched the education in social networking sitesproblems. Also was seen semantic web, expert system, cloud computing. Usual education may be also usual with using internet too, when it comes to online social networking, websites are commonly used. These websites are known as social sites. Social networking websites function like an online community of internet users. Depending on the website in question, many of these online community members share common interests in hobbies, religion, politics and alternative lifestyles. Once you are granted access to a social networking website you can begin to socialize. This socialization may include reading the profile pages of other members and possibly even contacting them. In this research work mixed three main part for create knowledge base which are cloud computing, expert system and finally semantic web. These three parts do this social networking site such unique.

References

1. The decree of the President of Republic Uzbekistan from December, 10th, 2012 № «About measures on the further perfection of system of learning of foreign languages» PP-1875. *http:www.lex.uz*
2. Al-Khalifa, H. S., and Davis, H. C. Harnessing the wisdom of crowds: How to semantically annotate web resource using folksonomies. In Proceedings of IADIS Web Applications and Research (WAR2006) (2006).
3. Alani, H., Dasmahapatra, S., O'hara, K., and Shadbolt, N. Ontocopi - using ontology based network analysis to identify communities of practice. IEEE Intelligent Systems 18 (2003), 18–25.
4. Andriessend, E., Soekijad, M., Veld, M. H. i. t., and Poot, J. J. Dynamics of knowledge sharing communities. Tech. rep., Delft University of Technology, 2001.
5. Angeli, C., Bonk, C., and Hara, N. Content analysis of online discussion in an applied educational psychology course. [On-line] (1998).
6. Archee, R. Using computer-mediated communication in an educational context: Educational outcomes and pedagogical lessons of computer conferencing. The Electronic Journal of Communication 3, 2 (1993). http://www.cios.org/getfile/ARCHEE_V3N293
7. Bales, R. F. Interaction Process Analysis. A Method for the Study of Small Groups. Cambridge, MA: Addison–Wesley, 1950/1951.
8. Balsamo, A. Technologies of the Gendered Body: Reading Cyborg Women. Duke University Press, New York, December 1995.
9. Bandura, A. Self-efficacy: Towards a unifying theory of behavioral change. Psychological Review 84 (1977), 191–215.
10. Berglund, A., and Eckerdal, A. What do cs students strive to learn? Computer Science Education 16, 3 (2006), 185–195.

11. Berkson, L. Problem-based learning: Have the expectations been met? Academic Medicine 68, 10 (1993), S79–S88.

12. Bolin, M. End-User Programming for the Web. PhD thesis, MIT, http://groups.csail.mit.edu/uid/ projects/chickenfoot/mbolin-thesis.pdf, May 2005.

13. Booth, S. Learning computer science and engineering in context. Computer Science Education 11, 3 (2001), 169–188.

14. Bordia, P. Face-to-face versus computer-mediated communication: A synthesis of the experimental literature. The Journal of Business Communication 34, 1 (1992), 99–120.

15. Borgatta, E. F., and Bales, R. F. Interaction of individuals in reconstituted groups. Sociometry 16, 4 (1953), 302–320.

16. Boulos, M., Maramba, I., and Wheeler, S. Wikis, blogs and podcasts: a new generation of web-based tools for virtual collaborative clinical practice and education. BMC Medical Education 6, 41 (August 2006).

17. Cayzer, S. Semantic blogging and decentralized knowledge management. Communications of the ACM 47, 12 (December 2004), 47–52.

18. Chin, M. Facebook statistics. http://static.ak. facebook.com/press/facebook_statistics.pdf, May 2007.

19. Citeulike. http://www.citeulike.org.

20. Classmates. http://www.classmate.com.

21. Cohoon, J. Recruiting and retaining women inundergraduate computing majors. SIGCSE Bulletin34, 2 (2002), 48–52.

[21] Collison, C. Connecting the new organisation. howbp amoco encourages post-merger collaboration.Knowledge Management Review (2000).
[22] Cooper, J., and Weaver, K. Gender andComputers: Understanding the digital divide.Lawrence Erlbaum & Associates, Mahwah, NJ, 2003.
[23] Dailymotion. http://www.dailymotion.com.
[24] del.icio.us. http://del.icio.us/.
[25] Engestr¨om, Y. Learning by Expanding: anactivity-theoretical approach to developmentalresearch. Orienta-Konsultit: Helsinki., 1987.
[26] Facebook. http://www.facebook.com/.
[27] Filmtrust. http://trust.mindswap.org/FilmTrust/.
[28] Flickr. http://www.flickr.com/.
[29] Friendsreunited.http://www.friendsreunited.co.uk.
[30] Friendster. http://www.friendster.com/.
[31] Garrison, R., Anderson, T., and Archer, W.Critical inquiry in a text-based environment:Computer conferencing in higher education.Unpublished manuscript, 2000.

[32] Goecks, J., and Cosley, D. Nuggetmine:Intelligent groupware for opportunistically sharinginformation nuggets. In Proceedings of the 2002International Conference on Intelligent UserInterfaces (IUI 2002) (2002), ACM, pp. 87–94.
[33] Goecks, J., and Mynatt, E. Leveraging socialnetworks for information sharing. ComputerSupported Cooperative Work 04 6, 3 (2004), 328–331.
[34] Gunawardena, C., and Zittle, F. Social presenceas a predictor of satisfaction within a computermediated conferencing environment. AmericanJournal of Distance Education, 11, 3 (1997), 8–26.
[35] Hamasaki, M., and Takeda, H. Find betterfriends? re-configuration of personal networks byneighborhood matchmaker method. In SWAFT(2003), pp. 73–76.
[36] Harasim, L. Shift happens: online education as anew paradigm in learning. Internet and HigherEducation 3 (2000), 41–61.
[37] Heath, C., and Heath, D. Finding just enough ofthat sticky stuff. Brandweek 48, 1 (2007), 21–25.
[38] Jackson, M. O. A survey of models of networkformation: Stability and efficiency. Group Formation in Economics: Networks, Clubs, and Coalitions(2003).
[39] Johnson, D. W., and Johnson, F. P. JoiningTogether. Group Theory and Group Skills. Boston:Allyn and Bacon, 1975/1997.
[40] Jordan, M. H., Feild, H. S., and Armenakis,A. A. The relationship of group process variables and team performance. a team-level analysis in a fieldsetting. Small Group Research 33, 1 (2002), 121–150.
[41] Kinnunen, P., and Malmi, L. Problems in problem-based learning - experiences, analysis and lessons learned on an introductory programming course. Informatics in Education 4, 2 (2005), 193–215.
[42] Kinnunen, P., and Malmi, L. Why students drop out cs1 courses? In Proceedings of ICER 2006 (2006), ACM, pp. 97–108.
[43] Kinnunen, P., McCartney, R., Murphy, L., and Thomas, L. Through the eyes of instructors: a phenomenographic investigation of student success. In Proceedings of ICER (2007).
[44] Kuehn, T. Communication innovation on a bbs: A content analysis. Interpersonal Computing and Technology: An Electronic Journal for the 21st Century 1, 2 (1993).
[45] Kuutti, K. Context and Consciousness: Activity Theory and Human Computer Interaction. Cambridge MIT Press, 1995, ch. Activity Theory as potential framework for human-computer interaction research, pp. 17–44.
[46] Lastfm. http://www.lastfm.com/.
[47] Laurillard, D. Rethinking University Teaching - 2nd Edition. Routledge-Falmer, London, 2002.
[48] Lave, J., and Wenger, E. Situated learning: Legitimate peripheral participation. 1-38, Palo Alto, CA, Institute for Research Learning, 1990.
[49] Leuf, B., and Cunningham, W. The Wiki Way. Addison Wesley, Boston, MA, 2001.
[50] Margolis, J., and Fisher, A. Unlocking the clubhouse: women in computing. The MIT Press, 2003.
[51] Marton, F., Dall'Alba, G., and Beaty, E. Conceptions of learning. International journal of educational research 19, 3 (1993), 277.
[52] Maybury, M., D'Amore, R., and House, D. Expert finding for collaborative virtual environments. Communications of the ACM 44, 12 (December 2001), 55–56.
[53] Mayfield, R. Social Network Dynamics and Participatory Politics. O'Reilly, Lulu.com, 2005, ch. 10, pp. 116–132.
[54] McDonald, D. W., and Ackerman, M. S. Expertise recommender: A flexible recommendation sytem and architecture. In Computer Supported Cooperative Work 2000 (2000), ACM, pp. 231–240.
[55] McPherson, M., Smith-Lovin, L., and Cook, J. M. Birds of a feather: Homophily in social networks. Annual Review of Sociology 27 (2001), 415–44.
[56] Mehrabian, A. Some referents and measures of nonverbal behavior. Behavior Research Methods and

Instrumentation 1, 6 (1969), 205–207.
12
[57] Microsoft live meeting. http://office.microsoft.com/en-us/livemeeting/default.aspx.
[58] Microsoft office live. http://www.officelive.com.
[59] Microsoft windows live. http://get.live.com/.
[60] Morales-Mann, E. T., and Kaitell, C. A. Problem-based learning in a new canadian curriculum. Journal of Advanced Nursing 33, 1 (2001), 13–19.
[61] Myspace. http://www.myspace.com/.
[62] Nathan, R. My Freshman Year: What a Professor learned by becoming a student. Cornell University Press, Ithaca, NY, August 2005.
[63] Nuutila, E., T¨orm¨a, S., and Malmi, L. Pbl and computer programming - the seven steps method with adaptations. Computer Science Education 15, 2 (2005), 123–142.
[64] Orkut. http://www.orkut.com/.
[65] Ounnas, A., Davis, H. C., and Millard, D. E. Towards semantic group formation. In The 7th IEEE International Conference on Advanced Learning Technologies (Niigata, Japan, July 2007), ICALT 2007. (in press).
[66] Ounnas, A., Liccardi, I., Davis, H. C., Millard, D., and White, S. A. Towards a semantic modeling of learners for social networks. In SW-EL (2006), Workshop at the AH 06, pp. 1–6. Dublin, Ireland.
[67] Preece, J. Online Communities - Designing Usability, Supporting Sociability. John Wiley and Sons Ltd, 2000.
[68] Prensky, M. Digital natives digital immigrants. On the Horizon, NCB University Press 9, 5 (October 2001).
[69] Prentice, D. A., Miller, D. T., and Lightdale, J. R. Asymmetries in attachments to groups and their members: Distinguishing between common-identity and common-bond groups. Personality and Social Psychology Bulletin 20, 5 (1994), 484–493.
[70] Reunion. http://www.reunion.com/.
[71] Rohde, M., Klamma, R., Jarke, M., and Wulf, V. Reality is our laboratory: communities of practice in applied computer science. Behaviour & Information Technology 26, 1 (January 2007), 81–94.
[72] Rourke, L., Anderson, T., Garrison, D. R., and Archer, W. Assessing social presence in asynchronous text-based computer conferencing. Journal of Distance Education/Revue de l'enseignement `a distance (2001).
[73] Saljo, R. Learning in educational settings: Methods of inquiry. P. Ramsden (Ed.), Improving Learning: New Perspectives (1988), 32–48, New York: Kogan Page.
[74] Sassenburg, K. Common bond and common identity groups on the internet: Attachment and normative behavior in on-topic and off-topic chats. Group Dynamics: Theory, Research, and Practice 6, 2 (2002), 27–37.
[75] Schellens, T., Van Keer, H., Valcke, M., and De Wever, B. Learning in asynchronous discussion groups: a multilevel approach to study the influence of student, group and task characteristics. Behavior and Information Technology 26, 1 (January 2007), 55–71.
[76] Schmidt, H. G. Problem-based learning: rationale and description. Medical Education 17, 1 (1983), 11–16.
[77] Schwen, T. M., and Hara, N. Community of practice: A metaphor for online design? The

Information Society 19 (2003), 257–270.
[78] Second life. http://www.secondlife.com/.
[79] Slattery, K. L. Loyalty, harm and duty: Pbl in a media ethics course. Public Relations Review 28 (2002), 185–190.
[80] S.R.Roxanne, and B.Wellman. Asynchronous learning networks as a virtual classroom. Communications of the ACM 40, 9 (September 1997), 44–49.
[81] Tajfel, H. Social Identity and Intergroup Relations. Cambridge University Press, Cambridge, 1982.
[82] Tinto, V. Classrooms as communities: Exploring the educational character of student persistence. Journal of Higher Education 68, 6 (Nov. - Dec. 1997), 599–623.
[83] Tinto, V. Learning better together: The impact of learning communities on student success in higher education. Journal of Institutional Research 9, 1 (2000), 48–53.
[84] Tuckman, B. W. Developmental sequence in small groups. Psychological Bulletin 63, 6 (1965), 384–399.
[85] Veoh. http://www.veoh.com/.
[86] Vivacqua, A., and Lieberman, H. Agents to assist in finding help. In CHI (2000), ACM, pp. 65–72.
[87] Vygotsky, L. S. Mind in Society. Cambridge: Harvard University Press, 1978.
[88] Wakeford, N. The embedding of local culture in global communication: independent internet cafes in london. New Media and Society 5, 3 (2003), 379–399.
[89] Walther, J. B. Computer-mediated communication: Impersonal, interpersonal, and hyper-personal interaction. Communication Research 23, 1 (1996), 3–43.
[90] Wegerif, R. The social dimension of asynchronous learning networks. Journal of ALN 2 (1998), 34–49.
[91] Wellman, B. Physical place and cyberplace: The rise of personalized networking. International Journal of Urban and Regional Research 25, Special Issue on "Networks, Class and Place" (2001). Edited by Talja Blokland and Mike Savage.
[92] Wenger, E. Communities of practice: Learning, meaning and identity. Cambridge University Press (1998).
[93] Wenger, E., and Snyder, W. M. Communities of practice: The organizational frontier. Harvard Business Review (2000).
[94] Wheelan, S. A., and Williams, T. Mapping dynamic interaction patterns in work groups. Small Group Research 34, 4 (2003), 443–467.
[95] Wiki choicetree. http://c2.com/cgi/wiki?WikiChoicetree, May 2005.
[96] Wikipedia

Websites:

http://gov.uz (Uzbekistan government web portal)
http://ru.wikipedia.org (free encyclopedia - Wikipedia)
http://google.com (Google search)

Atachments

#	North	East	weight of goods kg	Address or organization
1	69.297118	41.294838	0	Storehouse
2	69.277319	41.327978	40	Center 5-57/13
3	69.268912	41.302201	40	Afrosiyob str.4B Asrtech
4	69.262319	41.283181	60	Nukus str.32 Fashion House LALI
5	69.296780	41.295402	80	Fargona str.7, Gasthhaus
6	69.297295	41.361193	80	children's city hospital №1 3 Yangi Shakhar str.
7	69.267683	41.375026	80	Zenit factory Yukori Korakamish str.
8	69.209114	41.369656	40	School №234 Karakamish 2/4
9	69.211743	41.362707	100	Hospital of National Security Service Karakamish str.2
10	69.208052	41.352802	40	TTU (Polytechnic) campus, Almazar district
11	69.235809	41.325330	20	Chorsu Sakimchon str.
12	69.253012	41.323082	40	Museum of history of Uzbekistan Abay str.5
13	69.271986	41.325523	60	Embassy of Germany Sharaf Rashidov str.15
14	69.282844	41.336657	120	Hotel International (Intercontinental) Amir Temur 107
15	69.272029	41.312727	80	NSS of Uzbekistan Matbuotchilar str.9
16	69.289710	41.292045	40	Central city station Mirabad District
17	69.267222	41.280065	80	Maternity home №8 Kakhkhar Drive 7
18	69.207874	41.299047	40	Global Study Chilanzar-7, Лутфий str.6
19	69.195815	41.299934	60	City oncological city 1 Bogiston str.
20	69.191352	41.299176	80	City emergency service (ambulance) Foziltepa str.40
21	69.204935	41.322708	80	Kukcha mosque Mannon Uygur str.
22	69.179452	41.301022	40	Laboratory of JSC"UzLITIneftgaz" Ali Kushchi str.
23	69.283671	41.268418	60	United City Hospital #2 Sergeli district
24	69.274830	41.357401	60	Clinical Hospital 7 Gayrati str.24
25	69.297532	41.361605	80	Children's city hospital №1 3 Yangi Shakhar str.
26	69.273242	41.381138	40	Computer technologies college of Kalanovo H.Z. Yunusaad
	Total weight		1540	
	Damas payload		450	
	Quantity of Damas		4	

```
/*
 *
 */
packageorg.metavrp.algorithm;

importjava.util.ArrayList;
importorg.metavrp.Algorithm;
importorg.metavrp.Problem;
importorg.metavrp.algorithm.GA.Gene;
```

```
importorg.metavrp.algorithm.GA.operators.OperatorsAndParameters;

public class GeneticAlgorithm implements Algorithm{

   // The Genetic Algorithm's operators and parameters
privateOperatorsAndParametersoperatorsAndParameters;

   // The Problem definition
private Problem problem;

   /**
    * Constructor that sets the problem definition
    * @param problem
    */
   // TODO: Verify that the problem has a minimum of 2 customers, 1 depot and
1 vehicle
publicGeneticAlgorithm(Problem problem) {
this.problem = problem;
   }

   /**
    * Constructor that initializes the operators and parameters and the problem
    * @paramoperatorsAndParameters
    */
   // TODO: Verify that the problem has a minimum of 2 customers, 1 depot and
1 vehicle
publicGeneticAlgorithm(OperatorsAndParametersoperatorsAndParameters,
Problem problem) {
this.operatorsAndParameters = operatorsAndParameters;
this.problem = problem;
   }

   /*
    * Getters and Setters
    */

   /**
    * Gets the operators and parameters
    * @return
    */
publicOperatorsAndParametersgetOperatorsAndParameters() {
returnoperatorsAndParameters;
   }

   /**
```

```
     * Sets the GA's operators and parameters
     * @param op
     */
public void setOperatorsAndParameters(OperatorsAndParameters op) {
this.operatorsAndParameters = op;
    }


    /**
     * Returns a list of all the genes
     */
publicArrayList<Gene>getGenes() {
        // Create an array of all the possible genes (vehicles and customers)
ArrayList<Gene> genes = new ArrayList<Gene>(problem.getCustomers().size()
+ problem.getVehicles().size());

        // Add the customers
genes.addAll(problem.getCustomers());

        // Add the vehicles
genes.addAll(problem.getVehicles());

return genes;
    }

    /**
     * Returns a list of all the genes properly cloned
     * @return
     */
    // TODO: This method is necessary?
publicArrayList<Gene>getClonedGenes() {
return (ArrayList<Gene>)getGenes().clone();
    }

    /**
     * Get the number of genes that the GA will use
     * This is the number of costumers + the number of vehicles
     * @return
     */
publicintgetNrGenes(){
returngetNrCustomers() + getNrVehicles();
    }

    /**
     * Returns the number of customers, as defined on the problem
```

```
     * @return
     */
publicintgetNrCustomers() {
returnproblem.getCustomers().size();
    }

    /**
     * Returns the number of vehicles, as defined on the problem
     * @return
     */
publicintgetNrVehicles() {
returnproblem.getVehicles().size();
    }

    /**
     * Returns the number of depots, as defined on the problem
     * @return
     */
publicintgetNrDepots() {
returnproblem.getDepots().size();
    }

    /**
     * Returns the problem that this GA will try to solve
     * @return
     */
public Problem getProblem() {
return problem;    }}
```