Online Social Networks: Status and Trends

George Pallis, Demetrios Zeinalipour-Yazti, Marios D. Dikaiakos

Department of Computer Science
University of Cyprus
1678, Nicosia, Cyprus

{gpallis,dzeina,mdd}@cs.ucy.ac.cy

Abstract. The rapid proliferation of *Online Social Network (OSN)* sites has made a profound impact on the WWW, which tends to reshape its structure, design, and utility. Industry experts believe that OSNs create a potentially transformational change in consumer behavior and will bring a far-reaching impact on traditional industries of content, media, and communications. This chapter starts out by presenting the current status of OSNs through a taxonomy which delineates the spectrum of attributes that relate to these systems. It also presents an overall reference system architecture that aims at capturing the building blocks of prominent OSNs. Additionally, it provides a state-of-the-art survey of popular OSN systems, examining their architectural designs and business models. Finally, the chapter explores the future trends of OSN systems, presents significant research challenges and discusses their societal and business impact.

1 Introduction

With the emergence of Web 2.0, end-users are placed at the heart of various Web technologies, which tend to reshape the future of the WWW in terms of its structure, design, and utility. In this context, *Online Social Networks (OSNs)* are emerging as a new type of "killer application" on the Internet, which can be considered as a natural extension of Web applications that establishes and manages explicit relationships between users. Specifically, an OSN consists of users who communicate with each other in an online setting in diverse ways. Nowadays, we have been witnessing the rapid rise of a large variety of OSN sites, which publish user-generated or aggregated content, allow users to annotate published content with tags, reviews, comments and recommendations, and provide mechanisms that enable the establishment of user communities based on shared interests [2], [11].

The first well-known OSN site, called SixDegrees.com, was launched in 1997; its name originates from the six degrees of separation concept. Six degrees of separation is the theory that anyone can be connected to any other person through a chain of acquaintances that has no more than five intermediaries. Through SixDegrees.com, users could create their profiles, have a list of friends and contribute information to their community. Although this site attracted million of users, it could not evolve into a sustainable business and closed down in 2000. The founder of SixDegrees.com believes that it was ahead of its time. From 2003, we witnessed a revolution and uptake of OSN sites that established most of nowadays most popular OSN sites. This revolution has brought a dramatic shift on the business, the cultural and the research landscape of the WWW [11]. Figure 1, presents a timeline that shows the evolution of OSN sites during the last decade.

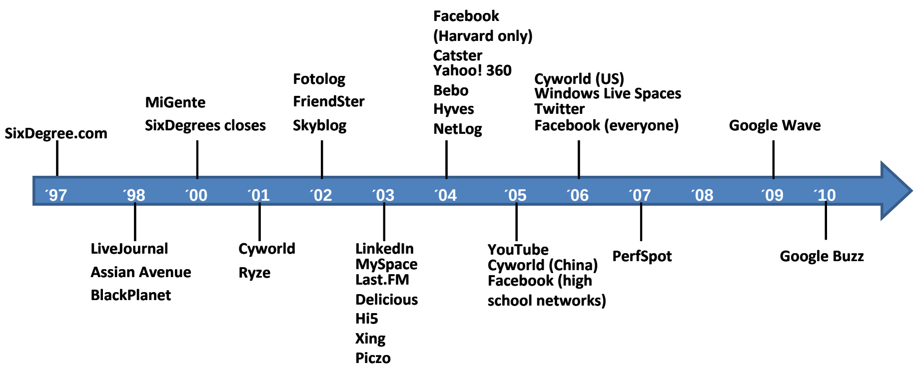

Fig1. Timeline of Online Social Network Sites.

According to Nielsen Online's latest research[1], social network and blogging sites are nowadays the fourth most popular activity on the Internet; this means that more than two-thirds of the global on-line population visit and participate in social networks and blogs. In fact, social media have pulled ahead of e-mail in the rank of the most popular online activities. Another interesting finding is that social networking and blogging accounts for nearly 10% of all time spent on the Internet. These statistics suggest that OSNs have become a fundamental part of the online experience on the WWW throughout the world.

The key breakthrough brought by OSN sites like Facebook, Myspace, Flickr, LinkedIn, and YouTube, and the main driving force behind their success, is that OSN sites promote the vision of a Human-centric Web, where the network of people and their interests become the primary source of information, which resides entirely on social networking services. Consequently, the main objective of OSN systems is to provide social networking functionality as a core service to a variety of high-level applications and services. In addition, online social networking opens new interesting problems and creates challenges for research in an environment that becomes increasingly complex, and less structured [1], [43]. Nowadays, OSNs have become the subject of numerous startup companies, offering users the ability to create, search and manage their own OSN communities.

---

[1]Nielsen Company: http://blog.nielsen.com/nielsenwire/wp-content/uploads/2009/03/nielsen_globalfaces_mar09.pdf

Currently, the main technical underpinnings of OSN infrastructures and services include Web 2.0 technologies, service-oriented software, caching, database and content distribution technologies. From a technical point of view, OSN sites provide APIs, software frameworks and open-source platforms that enable application developers to build applications and manipulate their content. The core of OSN platforms is the social graph, where nodes represent individual actors within a social network and edges represent the interdependencies between the actors, which is integrated with new applications [11]. Social networking services make the social graph an integral element of their backend infrastructure, and provide direct access to parts of the graph through their end-user interfaces. In addition, the advent of the Linked Open Data W3C project opens new perspectives to OSNs allowing user contributed content to become even more open and accessible. Specifically, Linked Open Data project[2] aims to connect data sets using semantic Web technologies such as RDF (Resource Description Framework) or RDFa (a simpler variation). RDF based descriptions of social data provide a rich typed graph and offer a much more powerful and significant way to represent online social networks than social graphs. In this context, several ontologies are used to represent social networks. For instance, FOAF[3] is used for describing people profiles, their relationships and their activities online and SIOC[4] ontology provides the basis for defining the users.

The goal of this chapter is to discuss OSN platforms from different perspectives. The main contributions of this chapter can be summarized as follows:

- We present a reference architecture for OSNs in order to establish some common terminology and for ease of exposition. Such an architecture facilitates the process of identifying the technical challenges that arise in constructing OSNs.

- We develop a comprehensive taxonomy of OSNs that provides an in-depth coverage of this field in terms of OSN organizational structure and service types. The main aim of our taxonomy is to explore the unique features of OSNs and to provide a basis for categorizing present and future development in this area.

- We present a state-of-the-art survey of prominent OSN platforms that provides a basis for instantiating the blocks of our taxonomy and for understanding the current social networking landscape. It also presents the underlying Web technologies that are currently exploited in the social networking field.

- We discuss significant open problems and research challenges that need to be addressed in order to develop efficient OSN infrastructures and services. Finally, we identify the strengths, weaknesses, and opportunities as well as the business and societal impact of OSNs.

[2] Linked Open Data project: http://linkeddata.org/
[3] The Friend of a Friend Project: http://www.foaf-project.org/
[4] SIOC-project: http://sioc-project.org/

The rest of the chapter is structured as follows: Section 2 describes the architecture of OSNs, providing an insight into the technological characteristics of these platforms. Section 3 presents the taxonomy of OSNs and Section 4 performs a detailed survey of prominent OSN platforms. Section 5 outlines the research challenges of OSNs, focusing on Decentralized OSNs and on the business and societal impact of this technology. Finally, Section 6 concludes this chapter.

2 Architecture of OSNs

An online social networking site is a Web site that:

- Acts as a hub for individuals to establish relationships with other persons (friends, colleagues, etc.). Each user articulates a list of other users with whom a connection is shared.

- Includes a wide range of tools for people to build a sense of community in an informal and voluntary way. Online users interact with each other, contribute information to the common information space, and participate in different interactive activities (e.g., photo uploading, tagging, etc.).

- Contains specific components that allow people to: define an online profile, list their connections (e.g., friends, colleagues), receive notifications on the activities of those connections, participate in group or community activities, and control permission, preference and privacy settings.

A reference architecture of OSNs is depicted in Figure 1. The entire system is formed by the following layers:

- **Data Storage layer**: This layer consists of two components: The *Storage Manager*, which is responsible for efficiently storing the information of social graphs and for handling increased database loads. This is usually achieved by adopting distributed memory caching. The other component, called *Data Store*, comprises the storage elements that store information items of a social networking service. Data Stores can be Multimedia Databases, User Profiles Databases etc.

- **Content Management layer**: This layer is responsible for three main tasks. Firstly, it facilitates the incorporation of social information from remote OSN sites through a Content Aggregator that gathers and organizes content from social media but also distributes it to other OSN platforms. Secondly, it facilitates the maintenance and the retrieval of the social content graph through the Data Manager. Thirdly, it controls the access of users by creating and maintaining an access control scheme.

- **Application layer**: Each OSN site supports numerous services such as search, news feeds, mobile access, etc. The services communicate with the data manager and the access control manager in order to analyze and manage the social content graph. The applications are provided to users through an application manager. The application manager facilitates the user interaction via a set of APIs. This component also includes a service framework for scalable cross-language services development. Such a framework allows users to deploy applications by abstracting the portions of each language that tend to require the most customization into a common library that is implemented in each programming language.

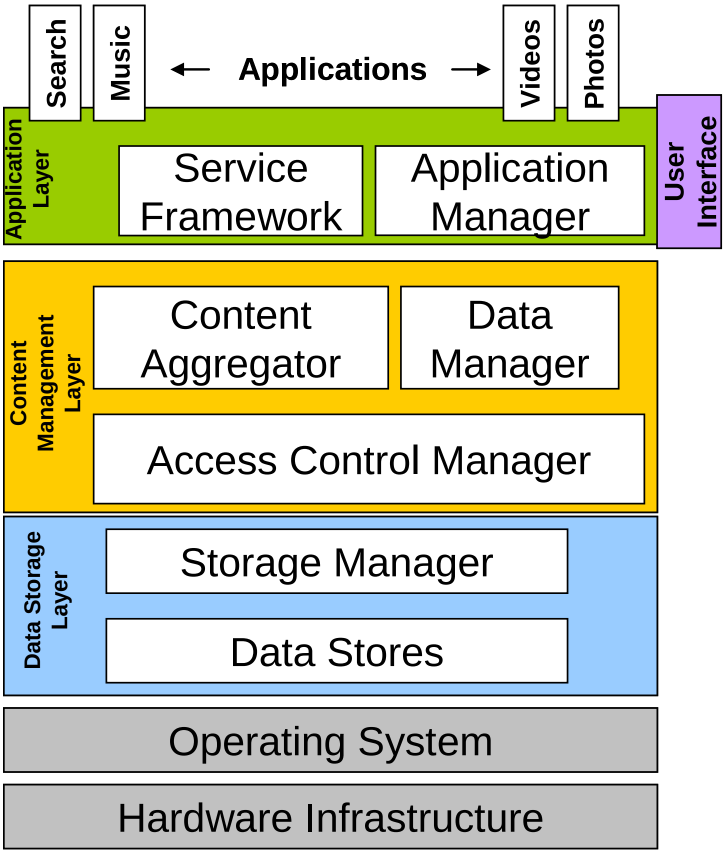

Fig 2. Reference Architecture of an Online Social Network Platform

The users interact with an OSN platform through HTTP requests. Each user can either be registered as an authorised user or be anonymous. Registered users usually have more rights than the anonymous ones (e.g., comment published articles, upload figures, etc.). The access control module is responsible for addressing privacy and security settings of OSN users. Specifically, most of the research in the field of OSN security and privacy has focused on the development of privacy-preserving techniques providing answers to issues such as the ownership of personal information and the protection of privacy [8], [15], [17], [38], [41].

Each OSN platform consists of multiple application servers, which provide a set of services and APIs. A user forwards a request to the OSN platform, which forwards it to the appropriate application server. A load balancer is responsible for monitoring the application servers of an OSN platform, balancing the load of requests, handling failover, and forwarding the requests to the application servers. The graph servers track and manage the connection relationships among users.

In terms of content distribution, cache servers speed up dynamic Web applications by alleviating the load of application servers. Specifically, OSNs involve a significantly different set of requirements compared to traditional Web applications. One of the major difference among OSNs platforms, are the relations and trust among users, which change over time. This behavior dictates the design of data placement, replication and distribution algorithms in OSNs. Another significant difference is that OSNs involve a large number of small files that need to be frequently accessed and updated by a large number of users and the propagation of file updates to guarantee data coherency. To further improve the system performance, OSN platforms distribute their content over Content Distribution Networks (CDNs) (e.g., Akamai, etc.) or Cloud Infrastructures. In a CDN setting [30], user requests are automatically routed to the nearest edge location, so content is delivered with the best possible performance. In a Cloud setting, the system is built over large clusters of processors. For instance, Facebook users can build their applications on Amazon Web Services (Amazon Elastic Compute Cloud (Amazon EC2) and Amazon CloudFront), improving reliability, flexibility, and cost-effectiveness.

Finally, the Data Stores of OSN platforms can be either centralized or distributed across multiple administrative domains. Centralized OSNs raise concerns regarding the protection of privacy and scalability. To overcome these limitations, data can be stored in a peer-to-peer infrastructure [12], [44], [45]. To support such a scheme, the reference architecture of Figure 1, should be extended by i) an overlay network layer on top of the operating system and network subsystem; and ii) an overlay management layer. The overlay network layer would provide mechanisms for node identity management, topology construction and maintenance, message routing, node search services, interfacing with local resources and the underlying fabric. The overlay management layer will provide mechanisms for authentication and authorization, decentralized monitoring, management, and adaptive control of peer-to-peer resources.

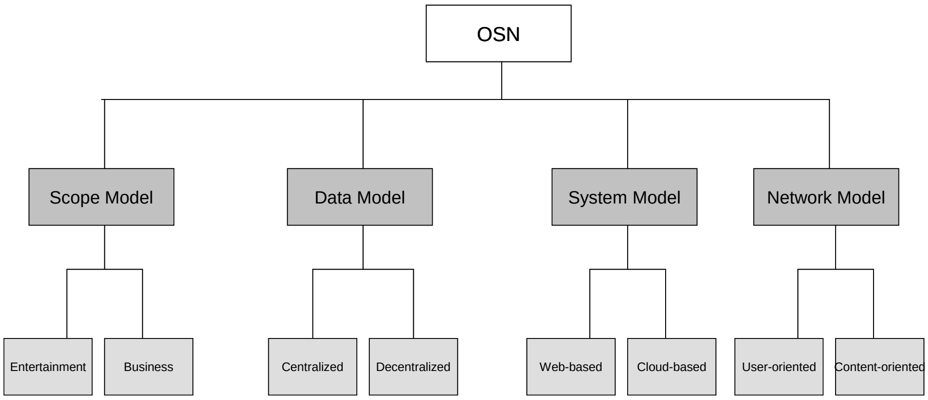

Fig. 2. OSN Taxonomy.

2 Taxonomy of OSNs

This section proposes a taxonomy that covers the key aspects of OSNs. This taxonomy is split into four branches. The first branch covers the scope of OSN systems in terms of activities. The second branch deals with the data model of OSNs, since the data model is the way in which data sources are stored in a system. The third branch, called system model, categorizes the OSNs regarding the hosting and content distribution of application servers. The fourth branch is from the point of view of the formation of users' network within OSN platforms. Figure 2 depicts the OSNs taxonomy.

In terms of their scope model, OSNs can be classified into the following two categories:

- **Entertainment**: Most OSNs are dedicated to entertainment. Their focus is on delivering fun and interactive social experience online to registered users. Popular OSN sites that are mainly entertainment-oriented are Facebook, Myspace, Hi5 and Flickr.

- **Business**: In this category, the focus of OSNs is to connect the world's professionals to make them more productive and successful. Through business OSNs, registered users create profiles that summarize their professional expertise and accomplishments. Indicative OSN sites in this category are LinkedIn and Xing.

The next dimension of our taxonomy is the data model followed by OSN providers. Here we identify the following categories:

- **Centralized:** the data of centralized OSNs are stored entirely in physical proximity (e.g., within a cluster or data center), concentrating the data of all their users under a single administrative domain. Today, most OSNs rely on centralized storage and functionality. However, centralized OSNs raise concerns regarding the protection of privacy and their scalability in the context of an expanding base of users and applications.

- **Decentralized:** the data of decentralized OSNs is distributed across multiple administrative domains [13], [35]. Application servers run on desktop machines (i.e., peers) owned by users. In general, hosting personal data on peers is more privacy-preserving than delegating control to a third-party service provider. In addition, this model is cheaper than acquiring dedicated centralized equipment. The main drawback of this approach is that peers might not be available continuously. Peers are prone to failures, reboots, power-offs, and network disconnections [12]. More discussion about the decentralized OSNs is given to section 5.

Our next dimension under discussion is the system model of OSNs. In particular, we identify the following two categories:

- **Web-based scheme:** the application servers are hosted by Web sites that provide a set of services and APIs. In such a scheme, the load balancer balances the load of requests, handles any failover, and forwards the requests to the appropriate application servers. In Web-based schemes, most OSN services are free to users.

- **Cloud-based scheme**: the application servers are hosted by a utility computing infrastructure such as Amazon Elastic Compute Cloud (EC2). In such a scheme, each user stores its own data on a personal virtual machine instance, called Virtual Server. The main advantages of this scheme are its high availability and its improved privacy since each user keeps its personal data in a Virtual Server residing in a Cloud computing environment. Also, Cloud-based schemes are usually integrated with CDN infrastructures (e.g., Amazon Cloud Front integrates with Amazon EC2). This results in distributing content to end users with low latency and high data transfer speeds. On the other hand, hosting data in a Cloud increases the costs due to utilizing a commercial infrastructure.

Finally, our taxonomy addresses the network model of OSNs that can be classified into the following two categories:

- **User-oriented OSNs**: The OSNs of this category emphasize on the social relationships. In such OSNs, content sharing is mainly among users who belong to the same community. Indicative OSNs of this category are the Facebook, MySpace and LinkedIn.

- **Content-oriented OSNs**: The users' network is not determined by the underlying social relationships but by their common interests. Indicative OSNs of this category are blog networks, question answering networks and video networks (e.g., YouTube).

4 Case Studies

This section presents various case studies of the most popular OSN platforms. For each OSN platform, we present a brief history, its scope, some technical issues and its business model. Table 1 summarizes the key characteristics of OSN platforms taking into account the taxonomy described in the previous section.

4.1 Facebook

History & Scope. Facebook (www.facebook.com) was founded by Mark Zuckerberg while he was a psychology student at Harvard University. In February 2004, Zuckerberg launched "The facebook"; the name was taken from the sheets of paper distributed to freshmen, profiling students and staff. Within 24 hours, 1200 Harvard students signed up, and after one month, over one half of the undergraduate population had a profile. In August 2005, it became Facebook.com and the address was purchased for $200,000. In September 2006, the network was extended beyond educational institutions to anyone with a registered email address. Nowadays, Facebook is one of the most popular online destinations with more than 300 million subscribers, whospend an average of 20 minutes on the site daily, contribute 4 billion pieces of information, 850 million photos, and 8 million videos every month. The main scope of Facebook is *entertainment*. Facebook is both a *user-oriented* and a *content-oriented* OSN site. Not only do the users create their networks among their friends, but also they can create networks based on common interests (e.g., Princeton alumni Network). Each registered user has a personal profile, adds friends and sends them messages, uploads photos, videos, links, and updates her personal profile to notify friends about herself.

Architectural Design. The data model architecture of Facebook *is centralized* and based on a typical hierarchical PHP Web application model with a layer of data caching. The caching layer is provided via the memcached open source software. Regarding the *application layer*, Facebook supports an Apache open source RPC mechanism called Thrift, which is used for both low-latency real-time RPC and persistent structured data storage across a variety of applications, such as Search, News Feed etc. The RPC language is influenced by CORBA's Interface Definition Language (IDL). Facebook supports various backend services that use the Hadoop, Scribe and Hive frameworks. Regarding the *data storage layer,* data (e.g., photos, statuses, and comments) are periodically stored to central repositories and relational (MySQL) databases. As far as the system model is concerned, Facebook application servers run on *Web-based* and *Cloud-based infrastructures* (i.e., Amazon EC2).

Business Model. Since the subscription to Facebook is free, the business model of Facebook depends on advertising. This business model is based on the principle that free subscriptions build audiences with distinct interests and expressed needs that advertisers will pay to reach. Facebook supports two advertising polices; the Pay for Clicks (also called cost per click or CPC) advertising allows advertisers to specify a certain amount that they are willing to pay each time a user actually clicks on their ad, whereas, the Pay for Views (also called cost per thousand impressions or CPM) advertising allows customers to specify how much they are willing to pay for 1000 views of their ad. Also, Facebook provides the option to target the ads to specific groups of users. This allows advertisers to write ad text that is more personalized, making their ad more appealing to the users they are reaching. Facebook supports several targeting options such as location, birthdays, likes and interests. An indicative success story of Facebook advertising is the following: "*Over 12 months, CM Photographics generated nearly $40,000 in revenue directly from a $600 advertising investment on Facebook. Of the Facebook users who were directed to CM Photographics' website from the ads, 60% became qualified leads and actively expressed interest in more information.*" The worth of Facebook's social network has been estimated to $15 billion[5]. It is remarkable to note that two-thirds of comScore's U.S. top 100 websites and half of comScore's Global top 100 websites have implemented *Facebook Connect*, a Facebook API that enables individuals and organizations to link their applications to Facebook. comScore is a global leader in measuring the digital world and the preferred source of digital marketing intelligence.

4.2 MySpace

History & Scope. MySpace (www.myspace.com) was developed in August 2003 by eUniverse employees who sought to mimic the more popular features of Friendster - a social networking website. The first MySpace users were eUniverse employees. In July 2005, MySpace was sold for US$580 million to Rupert Murdoch's News Corporation. MySpace became the most popular social networking site in the US in June 2006. Nowadays, MySpace is one of the fastest growing OSN sites with 300 million users. Its main scope is *entertainment*. MySpace is a *user-oriented* OSN site. Each registered user can share among friends a personal profile, photos, videos, links, etc. The user can also play games and update her personal profile to notify friends

[5] http://www.facebook.com/press/releases.php?p=8084

about herself. MySpace also runs event management, such as product launches and album launches for major labels which provides additional revenue streams.

Architectural Design. MySpace architecture is *centralized* and based on ASP.Net 2.0 Web application model with a layer of data caching and extracted services components. Specifically, MySpace consists of more than 500 Web servers running windows 2003/IIS 6.0/APS.NET. The *Data Stores* of MySpace consist of 1200 cache servers running 64-bit Windows 2003 with 16GB of objects cached in RAM, and 500 database servers running 64-bit Windows and SQL Server 2005. The caching layer is provided via the memcached open source software. Within the *data storage layer*, data is stored in central repositories and relational databases. As far as the system model is concerned, MySpace application servers run on *Web-based* and *Cloud-based infrastructures* (Amazon EC2).

Business Model. The subscription to MySpace is also free. Therefore, the business model of MySpace is also dependent on advertising. Similarly to Facebook, MySpace supports two advertising polices: the Pay for Clicks and the Pay for Views. With a rich set of demographic and behavioral data, MySpace enables advertisers to target their precise audience based on attributes, interests and activities. All the services are offered free to users. Users can create new content to attract, on their turn, new users. This has led several companies to advertise their products through the MySpace platform.

4.3 Hi5

History & Scope. Hi5 (hi5.com) was founded in 2003 by entrepreneur Ramu Yalamanchi, who is also its current Chief Executive Officer. According to comScore, in 2008 Hi5 was the third most popular social networking site in terms of monthly unique visitors. Nowadays, Hi5 is available in 50 languages and is one of the most popular online destinations for *entertainment* with more than over 60 million active users. Hi5 provides a robust platform for third-party developers to integrate games, content, and other applications. Hi5 is a *user-oriented* OSN site where each registered user can upload photos, videos, songs, personal information and share them with friends. Also, the users can join groups with semantically common interests.

Architectural Design. Hi5 architecture is based on an N-tiered Java architecture model. The *Data Stores* of Hi5 consists of PostgreSQL database servers. Hi5 runs using open-source software on a Linux platform. Specifically, Hi5 uses Linux Servers running SuSE Enterprise Linux, Apache and Lighttpd Web servers, the Squid proxy server for Web acceleration, Resin and Tomcat Java Application Servers, Struts for Model-View-Controller (MVC), Spring for Java Application Framework, iBatis for Object Relational Mapping, the Lucene library for indexing, and the Enunciate as Web service deployment framework. The caching layer is provided via the memcached open source software. Regarding the *data storage layer*, data is stored in a centralized configuration of the PostgreSQL object-relational database system. All queries are centrally maintained in XML files. Regarding the system model, application servers run on a *Web-based infrastructure*.

Business Model. Like previous OSNs, the business model of Hi5 also depends on advertising. Hi5 is collaborating with SponsorSelect - a premium advertising network that is reinventing behavioral targeting –in order to allow its users to choose the advertising they wish to see. By allowing users to self-select advertising, advertising is more relevant and performs better for advertisers, publishers (a publisher displays ads, text links, or product links on its Web site) make more money and can provide better content and services to users for free.

4.4 Flickr

History & Scope. Flickr (www.flickr.com) is the most popular OSN dedicated to photo and video sharing, where millions of photos are uploaded, tagged and organized by more than 8.5 million registered Web-users. Flickr was founded by Stewart Butterfield and Caterina Fake. In February 2004, Flickr was launched by Ludicorp, a Vancouver-based company. In March 2005, Flickr was acquired by Yahoo!. Flickr is a *user-oriented* OSN site and its main scope is *entertainment*. In Flickr, every user enters/selects new tags for a particular photo/video and the system suggests related tags to user, based on the tags that the user or other people have used in the past along with (some of) the tags already entered.

Architectural Design. Flickr architecture is based on a typical hierarchical PHP Web application model with a layer of data caching. The *Data Store* of Flickr consists of 62 MySQL databases across 124 servers, with about 800,000 user accounts per pair of servers. The MySQL databases are hosted on servers that are Linux-based, with a software platform that includes Apache, PHP, shards, Memcached, Squid, Perl and Java. The system administration tools include anglia for distributed system monitoring and Cvsup for distributing and updating collections of files across a network. Flickr supports tools for image processing (ImageMagick) and deployment (SystemImager). Data (e.g., photos, videos and tags) are stored in *central repositories* and relational MySQL databases. Regarding the system model of Flickr, application servers run on *Web-based* and *Cloud-based infrastructures* (Amazon EC2).

Business Model. With 3 billion pictures, Flickr is the biggest repository of digital images and videos on the Web. The basic account is for free, but Flickr charges users that want a professional and more sophisticated account. Flickr follows a consolidated business model, the “freemium” model, in which the registrations fees are varied from free to a premium "pro" version, with the latter featuring more capabilities.

4.5 LinkedIn

History & Scope. LinkedIn (www.linkedin.com) was founded in 2003 by Reid Hoffman. LinkedIn is an interconnected network of experienced professionals from around the world, representing 150 industries and 200 countries. In December 2009, LinkedIn had more than 55 million registered users and it was available in four languages. The scope of LinkedIn is mainly for *business*, allowing users to maintain a list of contact details of people they know and trust in business. Through this network people can find jobs and business opportunities, whereas employers can post and distribute job listings for potential candidates. LinkedIn is a *user-oriented* OSN site where registered users create their networks by sending personal invitations. A key feature of LinkedIn is that registered users can be recommended by someone in one's contact network.

Architectural Design. The data model of LinkedIn is *centralized.* It is an open architecture that consists of one monolithic Web application. Its *Data Store* includes a set of databases and a social network graph. LinkedIn runs on the Solaris operating system, uses Tomcat and Jetty as application servers, Oracle and MySQL as Databases, Spring for Java Application Framework, the Lucene library for indexing and ActiveMQ for Java Message Services (JMS). Web applications provide the GUI to the user and update the databases directly. Regarding the system model, application servers run on a *Web-based infrastructure*.

Business Model. LinkedIn is free to join. In addition, LinkedIn offers a premium version providing more tools for finding and reaching the “right” people. Specifically, with a premium account users can send messages directly to people and search for profiles that do not belong in their network. An indicative success story is when LinkedIn drove highly users to the MAZDA6 site and delivered some of the highest KPI ratings of all lifestyle sites on the plan.

4.6 Twitter

History & Scope. Twitter (www.twitter.com), founded by Jack Dorsey, Biz Stone and Evan Williams in March 2006 and launched publicly in July 2006, is a social networking and micro-blogging service that allows users to post their latest updates. An update is limited to 140 characters (called tweets) and can be posted through a Web form, a text message, or an instant message. Tweets delivered to the author's subscribers who are known as “followers”. Senders can restrict their posts to specific friends or, by default, allow open access. Registered users can also follow lists of authors instead of following individual authors. The scope of Twitter is twofold: *business* and *entertainment*. For instance, Twitter has been used for campaigning (2008 US presidential campaign), educational purposes, public relations etc. Twitter is a *content-oriented* OSN site since a user’s network is determined by the underlying social relationships; users create their networks by becoming “followers”.

Architectural Design. The data model of Twitter *is centralized* and based on the Ruby on Rails Web application framework with a layer of data caching. The caching layer is provided via the memcached open source software. Regarding the *data storage layer*, data is periodically stored to central repositories and relational (MySQL) databases. Also, it supports networked resource monitoring tool (Munin and Nagios) for analyzing resource trends**.** As far as the system model is concerned, Twitter application servers run on *Web-based infrastructures*.

Business Model. Twitter is free for all registered users. Contrary to most OSN sites, Twitter does not provide any advertising policy. In addition, it does not support any premium account. Nowadays, Twitter is in beta test with providing enterprise subscriptions that would target corporate customers. The idea to provide enterprise subscription is based on the assumption that the more businesses use Twitter, the more ways the company will find to monetize their traffic.

4.7 YouTube

History & Scope. YouTube (www.youtube.com), founded by Steve Chen, Chad Hurley and Jawed Harim in February 2005, is a social networking that allows users to post their videos. In November 2006, YouTube was bought by Google for $1.65 billion. Recently, YouTube has been ranked as the fourth most visited Website on the Internet According to comScore, YouTube is the dominant provider of online video in the US. It is estimated that 20 hours of new videos are uploaded to the site every minute. In March 2008, YouTube's bandwidth costs were estimated at approximately US$1 million a day. The scope of YouTube is for *entertainment*. All users can watch open videos, while registered users are permitted to upload an unlimited number of videos. YouTube is a *content-oriented* OSN site since users' network is determined by users' common interests.

Architectural Design. The data model of YouTube is *decentralized* and based on the distributed storage system, called BigTable[6]. BigTable developed at Google and its scope is to store efficiently large-scale structured data. To further improve its performance, BigTable is used with MapReduce[7], a framework for running parallel computations. Regarding the content delivery, most popular content is moved to a CDN provider (Akamai), whereas, less popular content is delivered through YouTube servers. A distributed multilevel cache is used for decreasing latency. NetScaler Web application controller is used for load balancing and caching static content. Regarding the *data storage layer*, data is periodically stored to MySQL databases. As far as the system model is concerned, YouTube application servers run on *Web-based* and *Cloud-based infrastructures* (Amazon EC2).

[6] BigTable: http://labs.google.com/papers/bigtable.html
[7] MapReduce: http://labs.google.com/papers/mapreduce.html

Business Model. YouTube is free for all the users (subscribers or not). The business model of YouTube depends on advertising (i.e., Google AdSense). Google AdSense uses its Internet search technology to serve advertisements based on Website content, the user's geographical location, users' tags, and other factors. Those wanting to advertise with Google's targeted advertisement system may enroll through AdWords. AdWords offers Pay for Clicks advertising, and site-targeted advertising. Thus, the business model of YouTube is based on mass collaboration. Providing free access to its users results to an increased number of users and, hence, to increased profits through increased advertisement rates and more advertisers.

Table 1. Main Characteristics of OSN sites.

OSN Platform	**Scope**		**Data Model**		**System Model**		**Network Model**	
	Entertainment	**Business**	**Centralized**	**Decentralized**	**Web-based**	**Cloud-based**	**User-oriented**	**Content-oriented**
Facebook	ν		ν		ν	ν	ν	ν
MySpace	ν		ν		ν	ν	ν	
Hi5	ν		ν		ν		ν	
Flickr	ν		ν		ν	ν	ν	
LinkedIn		ν	ν		ν		ν	
Twitter	ν	ν	ν		ν			ν
YouTube	ν			ν	ν	ν		ν

5 Future Research Challenges

Despite the fact that most OSN sites are centralized, nowadays, we observe a paradigm shift from centralized to distributed infrastructures [12], [35]. In general, decentralization can provide answers to issues that have raised controversy in the context of centralized OSNs, such as the ownership of personal information and the protection of privacy; problems in cross-platform service provision and user lock-in [15], [17]. Decentralization promises higher performance, fault-tolerance and scalability in the presence of an expanding base of users and applications. OSN decentralization has been identified as a key research challenge [12]. However, the transition to a fully distributed architecture so as to scale up OSNs is nontrivial. It is often a costly endeavor and specially challenging for OSNs that were not designed to be fully distributed from day one [12], [35]. In this context, this paradigm shift gives rise to many research questions intersecting networking, security, distributed systems and social network analysis, leading to a better understanding of how technology can support social interactions.

In the following sub-sections we present requirements and challenges towards developing efficient *Decentralized Online Social Networking (DOSN)* infrastructures and services. Finally, we conclude this section by presenting the business and social impact of OSNs.

5.1 Overlay Networking

Researchers have analyzed different properties of OSNs, mainly focusing on their formation and evolution as well as their information propagation over the network [1], [29], [34]. The advent of DOSNs creates new perspectives and challenges in networking [12], [43]. The objective is to build fundamental overlay services that will form the foundation for DOSN services and applications. This requires the investigation of novel architectures, algorithms and protocols for overlay networks of peers so as to provide the run-time environment and the basic communication functionality required by DOSN services and applications.

Therefore, self-management techniques, P2P publish/subscribe mechanisms and software component models should be developed. Thus, it is important to devise algorithms for self-management of the existing network infrastructure as well as for efficient group-based communications and information sharing (publish/subscribe) among the participants of DOSNs. In addition, a publish/subscribe solution must be able to effectively deal with large populations of dynamic users, large numbers of topics, arbitrary subscription patterns and be robust to malicious peer behaviour. Finally, adaptive component models for DOSNs will enable the underlying infrastructure to continuously operate under ever changing network conditions, enabling the safe manipulation and configuration of existing overlay resources at runtime.

Finally, it is also important to understand how the workload of DOSNs is reshaping the Internet traffic as this is valuable in designing the next-generation Internet infrastructure and content distribution systems (e.g., CDNs). Although current DOSNs contribute a lot less than peer-to-peer applications in terms of bytes, DOSNs might add features that increase the per-user bandwidth demand [12]. Given this potential of traffic explosion (e.g., when video becomes popular within a OSN site), it is crucial to explore the network-level dynamics of DOSNs [34].

5.2 Privacy & Trust

The amount of digital content today circulated in OSNs is enormous providing personal information about their users (i.e., profiles, friend relationships, daily activities, photos, videos, etc.). Although security and privacy concerns can prevent such efforts in practice, the problem remains since the data is located on a single server. Up to now, many researchers have started to work on improving the access control systems provided by OSNs. Specifically, most of the research in the field of OSN security and privacy has focused on the development of privacy-preserving techniques to mine OSN data [14]. In addition, several research efforts [8], [15], [17], [38], [41] focus on addressing the restrictions of protection mechanisms provided by current OSNs. Decentralized OSNs address many privacy concerns since the personal data of users is distributed across multiple administrative domains. This provides users more control over their content, reducing the system's vulnerability. Therefore, novel privacy-aware access control solutions should also be developed, providing users as much control as possible over their data and the way they are protected. In [35], the authors compare the privacy, cost, and availability tradeoffs for DOSN schemes.

Considering that OSNs host a variety of personal data, it is also crucial to identify possible threats arising from distributed malicious data and content in DOSNs. In the context of Web, it has been observed a recent increase of exploits, such as drive-by downloads [32] and malicious documents [23], or other related Web threats such as Cross-Site Scripting (XSS) [6], [10] and Cross-Site Request Forgery (CSRF) [9]. OSNs are considered as ideal targets for this kind of threats, since they are usually composed by complex AJAX interfaces that serve millions of users who trust each other. For instance, a social application may be used to launch DoS attacks against third parties [5] or a fake user account may be used for SPAM distribution. Regarding the DOSNs, data replication takes place in user's desktops. This further augments the probability of hosting malicious content in a user's computer machine. Therefore, a DOSN must provide the required mechanisms for the identification and further prevention of malicious content distribution. Thus, it is important to provide practical workarounds and directed guideline for detecting and preventing further distribution of malicious content, where this is possible. This can be achieved by applying data encryption policies. Such policies ensure that data are accessible to users who have the right private keys. In [42] a secret sharing protocol for DOSNs is presented.

5.3 Knowledge Discovery and Search

The role of network structure on the Web has grown in significance in the field of knowledge discovery and information retrieval, stimulated to a great extent by the importance of link analysis in the development of Web search techniques. But the Web has always contained a second network, less explicit but equally important, and this is the social network on its users, with latent person-to-person links encoding a variety of relationships including friendship, information exchange, and influence. Developments over the past few years - including the emergence of OSN systems and rich social media, as well as the availability of large-scale e-mail and instant messaging datasets - have highlighted the crucial role played by OSNs, and at the same time have made them much easier to uncover and analyze [1].

The study of evolution of OSNs and their properties have been one of the central areas of social network analysis. In particular, it is inherently related with the problem of predicting particular attributes of OSNs [18]. A large body of work has focused on the study of global evolution of networks and the identification of communities [7], [21]. Other recent research work has focused on developing algorithmic tools for the analysis of evolving networks [34], [37] as well as inferring users' profiles based on social graphs [27].

There is now a considerable opportunity to exploit the structure, evolution and information content inherent in DOSNs since the prospect of decentralization raises a number of interesting research challenges. However, the complexity of DOSNs requires devising novel algorithms and mechanisms in order to study and model the evolution of DOSNs in both macroscopic and microscopic level. At the macroscopic level mining patterns that characterize the evolution of the network when it is viewed as a global structure should be investigated. At a microscopic level algorithms that identify primitive "patterns" of evolution will be developed [21]. A specific prediction problem that has drawn an amount of interest in the research community is the link-prediction problem [24]. The question is to infer which new interactions among the members of a DOSN are likely to occur in the near future. In addition, it is crucial to study the problem of link formation in DOSNs. New models for the evolution and the formation of links should be built by integrating all available information: topological structure of the graph, content information of the users of the network, as well mined patterns of evolution. Considering that DOSNs are very large and complex, methodologies and tools from the field of Complex Networks should be used in order to exploit them [31]. Also, methodologies and mechanisms should be investigated that allow identifying interesting patterns in the network and create/select features for characterizing the various components of the system and predicting their evolution. Such features can be the popularity of certain items or the existence of a link between two items, and they can be used for designing improved ranking models, providing early characterization of spam, early characterization of being member of a community (strongly connected components), leader, reputation of authors, etc.

In terms of searching, the current evolution of OSN sites is characterized by an increasing availability of online services and novel search facilities (e.g., services for searching scientific literature, photos, videos, vacation offers, travels, restaurants, online shops, and so on). A recent survey [28] has shown that people turn to OSNs rather than typical search engines or Q&A sites for certain question types and topics. However, the quality of their answers goes much beyond what can be achieved via conventional, general purpose search engines. Authors in [16] describe a social model of user activities before, during, and after search. Recently, Horowittz and Kamvar [20] presented a large-scale social search engine, called Aardvark. Contrary to traditional Web search engines, the scope of Aardvark is not to find the documents that satisfy the user's information need, but to find the person that can satisfy the user's information need. This person belongs to the user's social network.

The shift in search paradigm that we observe in OSNs opens up a number of interesting research questions in information retrieval. The main challenge posed by content in DOSN sites is the fact that the distribution of quality has high variance: from very high-quality items to low-quality, sometimes abusive content [4], [13]. This makes the tasks of filtering and ranking in such systems more complex than in other domains [40]. Also, trust in a traditional search engine is based on authority, whereas in a social search engine, trust is based on intimacy. However, social media systems present inherent advantages over traditional collections of documents: their rich structure offers more available data than in other domains. In addition to document content and link structure, DOSNs exhibit a wide variety of user-to-document relation types, and user-to-user interactions [2], [3], [22], [39].

5.4 Business and Social Impact

Online social networking is a complex, large and rapidly expanding sector of the information economy whose impact is expected to be far-reaching. User-generated content is causing changes in the traditional content/media industry structure. In the future, community features will be an integral part of all digital experiences - from information/publishing to business and entertainment. Companies providing services for social networking and media (e.g., sysomos[8] - a product suite that provides customers with tools for measuring, monitoring, understanding and engaging with the social media landscape) or adding social networking features to existing services must anticipate significant growth.

From technological perspectives, the success of OSN sites represents a growing threat to the monopoly of Google. For instance, Facebook has become one of the most popular online platforms with more than 300 million registered users (one fifth of all Internet users – circa August 2009) who spend an average of 20 minutes on the site every day, contributing 4 billion pieces of information, 850 million photos, and 8 million videos every month. This means that Facebook not only has it become perhaps

[8] Sysomos: http://www.sysomos.com/

the largest source of personal data online, but also it has embarked on an ambitious effort to challenge Google's position as the dominant driver of Web traffic and the dominant power in Web advertising. This challenge has led to the introduction by Google of the OpenSocial API, which promises to provide functionality similar to that of Facebook's platform, even though relying on open-standard technologies. Also, to the recent introduction by Google of the Buzz social networking services, which is integrated with Gmail.

The key added ingredient of OSN platforms is their social dimension with the aim of linking users together to facilitate their interaction and make it richer and more productive. The power of people interacting with people in an online setting has driven the success or failure of many companies in the Internet space. Kees Winkel (http://futurecase.wordpress.com/) argues that 2% of nodes in a social network are all that need to be reached to ensure an idea or marketing initiative successfully. A white paper[9] from AT&T discusses the business impact of social networking. According to this paper, the impact of social networking in business is catalytic, driving several companies to change their vision and organization.

Except of business, OSNs have also social impact. Authors in [36] studied the impact of OSNs in marketplaces and observed that social networking improves e-commerce presenting a very positive impact. Information from OSNs can also be exploited to improve Internet search engines [26], while others have applied this information to increase profits from viral marketing [33]. OSNs have also applications in Vehicular ad hoc Networks (VANETs) where groups of vehicles' drivers socialize and communicate with each other in order to inform for the roads' conditions [19]. As far as the education is concerned, OSNs can be helpful to a students learning environment, as long as it is used correctly and responsibly [25].

6 Conclusion

OSNs are popular infrastructures for information sharing, communication and interaction on the Internet. With over half a billion users, OSNs are nowadays a mainstream research topic of interest for computer scientists, economists, sociologists etc.

In this chapter, we have analyzed and categorized the infrastructural and technical attributes of OSNs. We have developed a comprehensive taxonomy for OSNs based on their scope, data model, system model and network model. We have also provided a detailed survey of the most popular OSNs and identified the underlying Web technologies that are currently in use in social networking domain. In doing so, the readers can gain an insight into the technologies, services and business models that are currently followed in this field. It is also presented the system architecture for OSNs, where the main components of an OSN platform are described.

[9] http://www.bligoo.com/media/users/1/50369/files/Business - Social Networking Impact.pdf

However, and despite their impressive success, OSN services face significant challenges that need to be addressed in order to improve the end-user experience and to allow for a healthy market expansion in the future. Consequently, content distribution, scalability and privacy issues are gaining more attention in order to meet up the new technical and infrastructure requirements of the next generation OSNs. This has led to a paradigm shift from a centralized infrastructure to a decentralized one. Thus, a section of this chapter is devoted to discuss the open problems and research challenges for OSNs. Finally, we explore the societal and business impacts of social networking.

References

[1] Y. Ahn, S. Han, H. Kwak, S. Moon, H. Jeong. Analysis of Topological Characteristics of huge Online Social Networking Services. Proceedings of the 16th International Conference on World Wide Web (WWW 2007), Banff, Alberta, Canada, May 2007.

[2] S. Amer-Yahia, L. Lakshmanan, C. Yu. SocialScope: Enabling Information Discovery on Social Content Sites. Proceedings of the CIDR 2009, Asilomar, CA, USA, 2009.

[3] S. Amer-Yahia, M. Benedikt, L. Lakshmanan, J. Stoyanovic. Efficient Network-aware Search in Collaborative Tagging Sites. Proceedings of VLDB Endow., 1 (1): 710-721, 2008.

[4] C. Anderson. The Long Tail: Why the Future of Business Is Selling Less of More. Hyperion Publisher, Jul 2006.

[5] E. Athanasopoulos, A. Makridakis, S. Antonatos, D. Antoniades, S. Ioannidis, K. G. Anagnostakis, E. P. Markatos. Antisocial Networks: Turning a Social Network into a Botnet. In Proceedings of the 11th International Conference on Information Security, Springer-Verlag, 146-160, Taipei, Taiwan, Sep., 2008.

[6] E. Athanasopoulos, V. Pappas, E. P. Markatos. Code-Injection Attacks in Browsers Supporting Policies. Proceedings of the 3rd Workshop on Web 2.0 Security & Privacy (W2SP), Oakland, California, USA, May 2009.

[7] L. Backstrom, D. Huttenlocher, J. Kleinberg, X. Lan. Group Formation in Large Social Networks: Membership, Growth, and Evolution. Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data mining, 44-54, New York, NY, USA., Aug. 2006.

[8] R. Baden, A. Bender, N. Spring, B. Bhattacharjee, D. Starin. Persona: an Online Social Network with User-defined Privacy. Proceedings of the ACM SIGCOMM 2009 Conference on Data Communication, 135-146, Barcelona, Spain, Aug. 2009.

[9] A. Barth, C. Jackson, J. C. Mitchell. Robust Defenses for Cross-Site Request Forgery. Proceedings of the 15th ACM Conference on Computer and Communications Security (CCS 2008), Alexandria, USA, Oct. 2008.

[10] A. Barth, J. Weinberger, D. Song. Cross-Origin JavaScript Capability Leaks: Detection, Exploitation, and Defense. Proceedings of the 18th USENIX Security Symposium (USENIX Security 2009), Montreal, Canada, Aug. 2009.

[11] D. M. Boyd, N. B. Ellison. Social Network Sites: Definition, History, and Scholarship. Journal of Computer-Mediated Communication, 13(1), 2007.

[12] S. Buchegger, A. Datta. A Case for P2P Infrastructure for Social Networks - Opportunities and Challenges. Proceedings of the 6th International Conference on Wireless On-demand Network Systems and Services, Snowbird, Utah, USA, Feb. 2009.

[13] S. Buchegger, D. Schioberg, Le-Hung Vu, A. Datta. PeerSoN: P2P Social Networking – Early Experiences and Insights. Proceedings of the 2nd Workshop on Social Network Systems (SocialNets 2009), Nuremberg, Germany, Mar. 2009.

[14] B. Carminati, E. Ferrari. Access control and Privacy in Web-based Social Networks. Journal of Web Information Systems 4(4):395–415, 2008.

[15] B. Carminati, E. Ferrari, A. Perego. Enforcing Access Control in Web-based Social Networks. ACM Transactions on Information & System Security 13(1):6, 2009.

[16] B. M. Evans, E. H. Chi. Towards a Model of Understanding Social Search. Proceedings of 2008 ACM Conference on Computer Supported Cooperative Work (CSCW 2008), San Diego, California, USA, Nov. 2008.

[17] P.W. L. Fong, M. M. Anwar, Z. Zhao. A Privacy Preservation Model for Facebook-Style Social Network Systems. Proceedings of the 14th European Symposium on Research in Computer Security (ESORICS 2009), Saint-Malo, France, Sep. 2009.

[18] L. Guo, E. Tan, S. Chen, X. Zhang, Y. Zhao. Analyzing Patterns of User Content Generation in Online Social Networks. Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France, Jun. 2009.

[19] M. Gupte, M.T. Hajiaghayi, L. Han, L. Iftode, P. Shankar, R. M. Ursu. News Posting by Strategic Users in a Social Network. Proceedings of the 5th Workshop on Internet and Network Economics (WINE'09), Rome, Italy, Dec. 2009.

[20] D. Horowittz, S. D. Kamvar. The Anatomy of a Large-scale Social Search Engine. Proceedings of the 19th International Conference on World Wide Web (WWW2010), Raleigh, North Carolina, USA, Apr. 2010.

[21] J. Leskovec, L. Backstrom, R. Kumar, A. Tomkins. Microscopic Evolution of Social Networks. Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data mining, Las Vegas, USA, Aug. 2008.

[22] J. Leskovec, D. Hattenlocher, J. Kleinberg. Predicting Positive and Negative Links in Online Social Networks. Proceedings of the 19th International Conference on World Wide Web (WWW2010), Raleigh, North Carolina, USA, Apr. 2010.

[23] W-J. Li, S. Stolfo, A. Stavrou, E. Androulaki, A. D. Keromytis. A Study of Malcode-Bearing Documents. Proceedings of the 4th International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment, 2007.

[24] D. Liben-Nowell, J. Kleinberg. The Link Prediction Problem for Social Networks. Proceedings of the 12th International Conference on Information and knowledge management (CIKM), New Orleans, Louisiana, USA, Nov. 2003.

[25] I. Liccardi, A. Ounnas, R. Pau, E. Massey, P. Kinnunen, S. Lewthwaite, M-A. Midy, C. Sarkar. The role of Social Networks in Students' Learning Experiences. In Working Group Reports on ITiCSE on Innovation and Technology in Computer Science Education. Dundee, Scotland, Dec. 2007.

[26] A. Mislove, K. P. Gummadi, P. Druschel. Exploiting Social Networks for Internet Search. Proceedings of the 5th Workshop on Hot Topics in Networks (HotNets'06), Irvine, CA, USA, Nov. 2006.

[27] A. Mislove, B. Viswanath, K. P. Gummadi, P. Druschel. You are Who you Know: Inferring User Profiles in Online Social Networks. Proceedings of the 3rd ACM International Conference of Web Search and Data Mining (WSDM'10), New York, Feb. 2010.

[28] M. R. Morris, J. Teevan, and K. Panovich. What do people ask their social networks, and why? A Survey study of status message Q&A behavior. Proceedings of CHI 2010, Atlanta, Usa, Apr. 2010.

[29] A. Nazir, S. Raza, D. Gupta, C.-N. Chuah, B. Krishnamurthy Network Level Footprints of Facebook Applications. Proceedings of Internet Measurement Conference (IMC 2009), Chicago, Illinois, USA, Nov. 2009 .

[30] G. Pallis, A. Vakali. Insight and Perspectives for Content Delivery Networks. Communications of ACM 49(1): 101-106, 2006.

[31] F. Papadopoulos, D. Krioukov, M. Boguna, A. Vahdat. Greedy Forwarding in Dynamic Scale-Free Networks Embedded in Hyperbolic Metric Spaces. Proceedings of IEEE INFOCOM, San Diego, CA, USA, Mar. 2010.

[32] N. Provos, P. Mavrommatis, M. Abu Rajab, F. Monrose. All your iFRAMEs point to Us. Proceedings of the 17th USENIX Security Symposium, San Jose, California, USA, Jul. 2008.

[33] M. Richardson, P. Domingos. Mining Knowledge-sharing Sites for Viral Marketing. Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Edmonton, Alberta, Canada, Jul. 2002.

[34] F. Schneider, A. Feldmann, B. Krishnamurthy, W. Willinger. Understanding Online Social Network Usage from a Network Perspective Proceedings of Internet Measurement Conference (IMC 2009), Chicago, Illinois, USA, Nov. 2009 ..

[35] A. Shakimov, A. Varshavsky, L. P. Cox, R. Caceres. Privacy, Cost, and Availability Tradeoffs in Decentralized OSNs. Proceedings of the 2nd ACM Workshop on Online Social Networks (WOSN 2009), Barcelona, Spain, Aug. 2009.

[36] G. Swamynathan C. Wilson, B. Boe, K. Almeroth, B. Y. Zhao. Do Social Networks improve e-commerce?: a Study on Social Marketplaces. Proceedings of the 1st Workshop on online Social Networks, Seattle, WA, USA, Aug. 2008.

[37] C. Tantipathananandh, T. Berger-Wolf, T., D. Kempe. A framework for Community Identification in Dynamic Social Networks. Proceedings of the 13th ACM SIGKDD International conference on Knowledge Discovery and Data mining, 717-726, New York, NY, USA. Aug. 2007.

[38] A. Tootoonchian, K.K. Gollu, S. Saroiu, Y. Ganjali, A. Wolman. Lockr: Social Access Control for Web 2.0. Proceedings of the 1st Workshop on Online Social Networks (WOSN 2008), Seattle, WA, USA, Aug. 2008.

[39] A. Ukkonen, C. Castillo, D. Donato, A. Gionis. Searching the Wikipedia with Contextual Information. Proceedings of the 17th International Conference on Information and Knowledge Management (CIKM), Napa Valley, California, USA, Nov. 2008

[40] M. V. Vieira, B. Fonseca, R. Damazio, P. Golgher, B. Davi, B. Ribeiro-Neto. Efficient Search Ranking in Social Networks. Proceedings of the 16th International Conference on Information and Knowledge Management, 563-572, Lisbon, Portugal, Nov. 2007.

[41] W. Villegas, B. Ali, M. Maheswaran. An Access Control Scheme for Protecting Personal Data. Proceedings of the 6th Annual Conference on Privacy, Security and Trust (PST 2008), Fredericton, New Brunswick, Canada, Oct. 2008.

[42] L. H. Vu, K. Aberer, S. Buchegger, A. Datta. Enabling Secure Secret Sharing in Distributed Online Social Networks. Proceedings of Annual Computer Security Applications Conference (ACSAC) 2009, Hawaii, USA. Dec. 2009.

[43] W. Willinger, R. Rejaie, M. Torkjazi, M. Valafar, M. Maggioni Research on Online Social Networks: Time to Face the Real Challenges. Proceedings of the 2nd Workshop on Hot Topics in Measurement and Modeling of Computer, 2009.

[44] D. Zeinalipour-Yazti, V. Kalogeraki, D. Gunopulos. Exploiting Locality for Scalable Information Retrieval in Peer-to-Peer Systems. Information Systems (InfoSys), Elsevier Publications, Volume 30(4):277-298, 2005.

[45] D. Zeinalipour-Yazti, V. Kalogeraki, D. Gunopulos. pFusion: An Architecture for Internet-Scale Content-Based Search and Retrieval. IEEE Transactions on Parallel and Distributed Systems (TPDS), 18(6): 804-817, 2007.