18

Online Social Networks and Social Network Services: A Technical Survey

Huangmao Quan
Temple University

Jie Wu
Temple University

Yuan Shi
Temple University

18.1 Introduction

Social functions are natural consequences of human societies. Before communication technologies, social functions tend to evolve within cultural boundaries, such as location and families. Communication technologies, from mountain top signaling to Voice-over-IP, have broken those boundaries more or less and enabled multi-culture social functions.

Empowered by low-cost, high-power personal computing devices, the combined computing and networking capabilities have created a fertile ground for innovative forms of social activities. *Online social network* (OSN) serves as a means of social activity and has become a mainstream information media in the industry and in the public. Both government and entrepreneurs recognized the value of OSNs and have put forth

efforts to capitalize on them. On the other hand, social networks also appear in public discourse. One significant example is social network analysis [1–3].

This article surveys OSNs. We provide our commentary mainly using technology factors, as well as their psychological and social backgrounds. Our article begins with a
Q1 background of taxonomy of OSNs by comparing similar counterparts. Our survey attempts to explain OSN II in Section 4, and proposes a taxonomy and retrospectively compares it with other online communities, and even pre-internet social networking communities. In Sections 3–5, we attempt to plot the possible evolution directions and future technology challenges. In general, this article states and attempts to answer the following questions:

- What are OSNs in comparison with conventional Web services?
- What different forms of infrastructure and application services are available in OSNs?
- Who are the main players in OSNs and what are they doing?
- What are the current developing trends and how successful will they be?
- What new services and technologies of OSNs are expected to appear in the future?

18.2 Background

It took the radio 38 years and TV 13 years to reach 50 million users. However, Facebook has added 100 million users in less than 9 months* and OSNs have become a part of our lives and changed us. On the other hand, we have changed OSNs as well. From its early form, which simply provides identity and relationship services to hundreds of services
Q2 that associate various applications with personal data. Social network services (SNSs) have gained their popularity and serve more and more granular target markets, such as medication, science, education, and so on. However, despite there being extensive amount of work on research of either online communities or off line social networks, they are generally not applicable in the context of OSNs. To make things even worse, with regard to the meaning of social network, sociology and computer science is lost in translation. In this section, our survey will reflect the development of OSNs in the past and then explicit the definition of OSNs in application and platform levels.

18.2.1 History

Generally speaking, in information technology, an OSN is basically a type of website that provides social identity and social relationship services: who are you, who do you connect with. To broaden our conceptual view, social networking, the concept of social phenomena per se, emerged before the Internet. Retrospectively, the first virtual community without propinquity, which we prefer to be called "off-line social network," appeared in the 17th century: the Royal Society of London formed a community through letter exchanging. Since then, various virtual communities became less and less

* www.socialnomics.net.

geographically binding and more and more based on common interests and activities [4]. Moreover, social networking*, or relationship initiation [5], as a common phenomenon for human beings in many aspects, originally existed in online virtual communities [6] before the appearance of OSNs. For example, dating sites and community sites supported lists of friends. Although most of those websites help strangers connect based on shared interests or activities—networking with others—it was not until the turn of the last century that new type of websites became recognizable, which extents and maintains pre-existing social networks by encouraging users to a create profile and affiliate with friends. They are called OSNs or social network sites interchangeably. Since the turn of the last century, some famous OSNs, such as Myspace and Facebook, were growing in popularity and proliferating. And more importantly, all the trends of human society have driven the growth of OSNs: internet capacities, hardware and software features, mobile communication, business model of Web 2.0, and so on. As a result, OSNs did not only hit the mainstream, but also became a global phenomenon [5].

18.2.2 What are OSNs?

Today, OSNs are used extensively as public social interactive and collaboration tools. The OSN distinguished itself in structure and behavior patterns from other relationship-initialized information systems, such as business relationship management system and collaboration software. According to Weyer's definition [7], interaction with OSNs is "an autonomous form of coordination of interactions whose essence is the trust cooperation of autonomous, but interdependent agents who cooperate for a limited time, considering their partners" interests, because they can thus fulfill their individual goals better than through non-coordinated activities'. Based on the definition, OSNs have four notable characters:

- No propinquity;
- No persistent connection;
- Trust based on interdependence;
- Autonomous collaboration.

No propinquity means an actor in OSN has no or little knowledge about the other actor in the other end of the tier (Latent or Weak ties [8]). No persistent connection means actors keep only temporary connections, unlike relationships in the real world, which are more stable. Trust, based on interdependent and autonomous collaboration means actors on OSNs have no obligation to serve others, their motivation of collaboration is from the awareness of interdependence.

Despite us believing that the above characters should be representative of most OSNs, we would also point out that the diversity of OSNs may fuzzy those characters. Some OSNs actually focus on strong ties, such as LinkedIn, which serves close communities that share more real-world connections. Whereas, OSNs vary widely by application and their key technological features are fairly consistent [5]. For example, most OSNs allow

* Social networking is activity initialed [7]. It a type of activity. While social network is an abstract Web connecting people by relationships.

people to articulate friends and publicly display connections, and their own profiles. Additionally, research in various fields developed its own taxonomies. Therefore, for better internal consistency, it is elastic to set a hard line between OSNs and general social networks. In our paper, OSNs are defined as Web-based services that allow individuals to:

- Construct a public or semi-public profiles within a bounded system,
- Articulate a list of other users with whom they share a connection,
- View and traverse their list of connections and those made by others within the system. The nature and nomenclature of these connections may vary from one site to other [5],
- Provide users' online presence to describe their current state and activities [9].

18.2.3 What is SNS?

Except structure and behavior patterns, OSNs also distinguish themselves from traditional Websites by providing various SNSs. Generally speaking, there are two types of SNSs. The first type is "organic" SNS, which is people-focused and embeds social network features within. For example, Twitter provides Microblogging [10], which has a core value of connecting friends and transferring ideas throughout a group of people. The second type is "hybrid" SNS, which is content-focused and combines traditional Internet services and social networking by integrating social features. For example, Flickr–Yahoo's photo-sharing Website combines photo repository and social networking features together. The first type aims towards the maintenance of pre-existing social networks and helps connect people based on common language, while the second type caters to diverse audiences.

All the functions of OSNs are delivered by SNSs. Some of them provides fundamental infrastructures that allow other services to build onto. We call them infrastructure SNSs. Another category only serves a specific purpose or application. We call them application level services. Despite that the SNSs may vary in their forms, they share the same goal, which is to fulfill human needs (Figure 18.1). It is human nature that drives us to socially connect with each other [11].

18.3 Social Network Services

While working on this survey, we found a dilemma while presenting all OSNs in a uniform classification. This is partly because there are two concepts—SNSs and social network platforms (OSNPs). These two concepts are logically hierarchic as applications and platforms, but normally used interchangeably. For example, despite most users agreeing that Twitter is an OSN, Twitter may not be a proper name for an SNS; microblogging would be the more appropriate one. Therefore, we intend to use OSNPs to describe Websites that host SNSs. To further clarify concepts of the OSN, this section will focus on SNSs, and the next section will talk about OSNPs.

In our paper, the term, SNSs, means Internet services provided by OSNs to *end-users*. Similar to other information systems, an OSN congregates a set of services, such as

	Expressing identity	Status & self-esteem	Giving & getting help	Affiliation and belonging	Sense of community
Blogs	✔	✔	✔		
Video, content sharing, tagging sites (e.g., YouTube, deli.ciou.us)	✔				
Self-forming groups (e.g., Yahoo or Google groups				✔	✔
Profile-driven social networks (e.g., MySpace, Linkedln, Facebook)	✔	✔		✔	
Rating, review sites (e.g., epinions, TripAdvisor)		✔	✔		
Purpose-driven social network (e.g., SparkPeople, Slashdot, Serma, Communispace)	✔	✔	✔	✔	✔

FIGURE 18.1 Human needs vs. social network services.

email, instant messaging, multimedia sharing, and so on. However, compared with traditional online services, the services provided by OSNs are more user-driven: giving more social context of users, SNS is usually described as an individual-centered service compared with traditional online communities where services are content-centered or group-centered [12]. Moreover, as we discussed in Section 2, SNSs may be "hybrid" products of traditional services. Because of the above reasons, we start by grouping SNSs according to their purposes. Then, we propose a map to explicit a taxonomy. We go on to survey specific examples of SNSs according to their classifications (Figure 18.2).

Broadly speaking, SNSs can be classified into infrastructure services and application services (Figure 18.3). Among them, infrastructure services provide the most basic and essential information about a social actor's identity, personal information, and his/her relationship connections.

- *Social profile*: Social actors' personal information of characteristics, such as name, gender, age, and so on.
- *Social identification*: The unique proof or evidence of identity, which is usable by other SNSs.
- *Social graph*: A relationship graph mapping of actors' friends and how they are related.

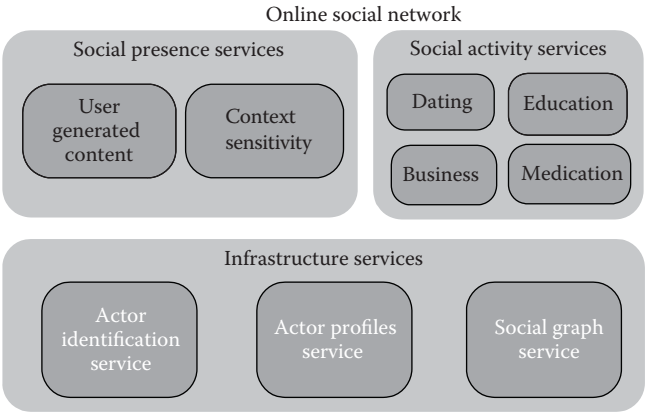

FIGURE 18.2 Classification of social network services.

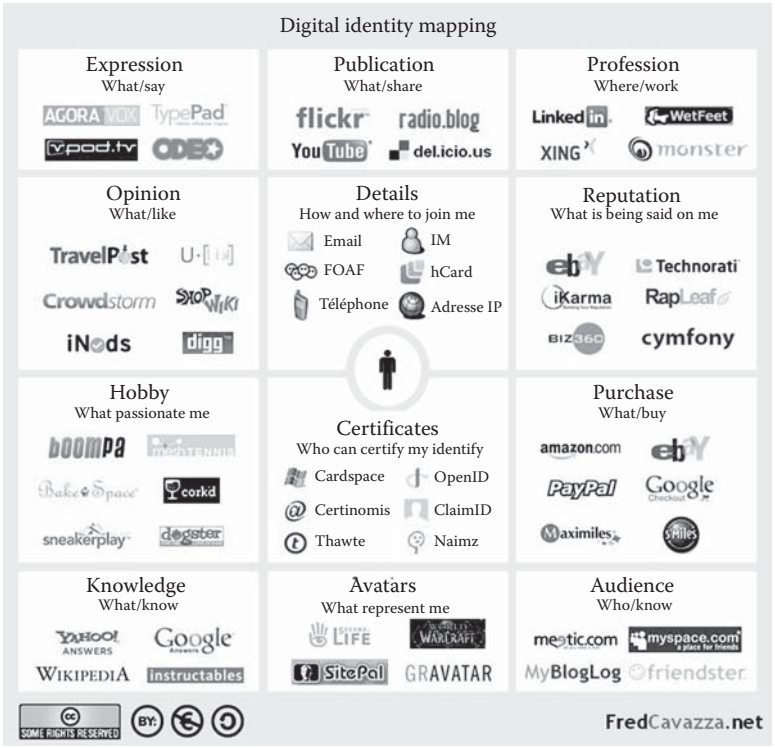

FIGURE 18.3 Example of social network services.

Supported by infrastructure services, OSNs provide various applications to facilitate social interactions and impress other people by social presence [13]. The term, "social presence", was derived from social present theory [14]. It is originally used to present awareness of the other person in a communication interaction. The interaction is effective if only social presence provides a meaningful indication of one person [15]. Thus, application level services can be classified according to their purposes:

- *Social interaction*: The services which include online communication of any sort, such as comment, instant message, and feedback.
- *Social presence*: Personal stamps created by users for others to indicate their state, log of activities, and performance.

Features of SNSs are different to different people depending on their activities. Therefore, our survey will concentrate efforts mostly in discussing services of infrastructure and social presence. In the next section, we will overview each SNSs, one by one.

18.3.1 Identification and Profile Service

The identity service, as one of the core components of OSNs, is critical to users because it affects others' awareness as well as bolsters users' self-esteem, sense of belonging, role, and hierarchy within online communities [16]. Additionally, persistent identity is required to build a stable connection of friends.

Identities are a set of characteristics that separate self from others. Although most OSNs require users to represent themselves authentically [5], most OSNs provide loose identity, for example, in twitter, no identification is required. Loose identity may not reflect users' authentic personality and lead to identity theft or impersonation (see Section 5 for details). In spite of coarse nature, in most conditions, identities of OSNs are reliable because other users can refer signals of profiles and public friend lists to validate them [5]. Moreover, different OSNs may select different strategies to rigidify their users. Some of them even encourage users to articulate fake identities or avatars [17].

The profile service of OSNs presents individual's choosing identity information. It can be updated in a specific time-frame and with a particular understanding of audience.

Firstly, the profile service is responsible for creating and maintaining profiles. For example, most OSNs ask users to fill out forms with a series of questions with regard to their personal information, which normally includes descriptors, such as name, sex, age, and interests. The profile is generated by the answers to those questions. Most OSNs also allow uploading profile photos.

Secondly, the profile service controls the visibility of a profile–profile closure. It provides an individual's fragmented profile according to audiences' discretion. For example, by default, basic profiles on LinkedIn and facebook are visible to anyone and even crawled by search engines. Therefore, they are publicly visible regardless of whether or not the audience has a connection with the user. Alternatively, the full profile, which includes sensitive information, such as marriage status and religious views, are available only to either paid accounts (LinkedIn) or connected friends (Facebook). Facebook and Myspace also allow users to choose whether they want their profile to be public or "Friends only." Facebook implied more complex settings for profile closure, for example,

users can grant permission to a certain network [18]. The counteraction of profile fragmentation is studied in Backstrom et al. [19] and Liu and Terzi [20], as a technical challenge. This will be discussed in Section 5.

Finally, based on this current business model [21], most OSNs' profile services have little or no interoperability with one another. Our paper will discuss this issue in detail in Section 5.

18.3.2 Social Graph Service

Social Graph Service is responsible for building, maintaining, and retrieving ties based on shared affinities. For SGSs; relationships, reputations, and searchs are three key components of the SGS: How to explicit connections—not only their connection status, but also how to quantitatively describe users feelings about connections, such as Ebay's reputation system.

For SGSs, one typical usage scenario is searching and extending connected people who share affinities or complementary capabilities [22]. For example, Alice and Bob belong to the same OSN and both are interested in politics. They are familiar with one another and trust each other. Alice has to work on a book that requires illustrations, and is looking for a freelance artist. By searching with SGS, Alice could be able to find Bob who is connected to Sarah, a student of art. Alice could then approach Sara for the project by sending a connecting request. Since Sarah knows that Jim and Bob trust each other and Bob trusts Sarah, it could be infered that possible trust between Alice and Sarah is high. Then, Sarah would confirm Alice's request with no hesitation. Finally, they will complete a happy and safe business transaction.

1. *Connection component:* Considering that a sense of affiliation is not equivalent to a true sense of community, connections in SGSs can either be bidirectional or unidirectional [5] to present relationships toward either individuals or groups, such as "became a fan" on facebook. To broaden our view, Brzozowski et al. [23] describes multiple relationship types in online communities. Contrary to popular belief, this research distinguishes allies and enemies in types of connections semantically. They argued that a better social network can be achieved by employing multi-typed connections.

   Compared with off-line social networks, people mostly use OSNs to reveal hidden relationships, which results in connecting people within their extended social networks [8]. For example: Bob and Alice know each other in reality. In other words, they share some offline connections. Alice has a friend, Ted, who works at BIG company as HR. All three have profiles on Facebook. Presumably, Bob needs to find a job. By tracking Alice's friends list on Facebook, Bob will have a good chance of tracking down Ted and contacting him with regards to a job. In this case, Facebook plays a bigger role more than just being an information media exchanging Website. It also extends and maintains relationships to benefit Bob's social activities—seeking a job in our example.
2. *Reputation component:* To avoid issues with regards to divergent reputation definition [7], in this paper, we simply define reputation as general opinions toward

individuals. There is extensive research on mathematical framework [24–27] for modeling caculatable reputations and the way reputations propagate [28].

Despite inferring reputations in OSNs being theoretically possible, applications that infer affinity and trusted third parties are limited. This greatly roots in structures of OSNs, which are not fully compatible with trust metrics. For example, Eigen Trust [29], a variation of the PagRank algorithm [30], provides a globally accepted trust rating as reputation. It was originally designed for P2P systems without considering the limitation of OSNs [31]. Despite lots of metrics trying to combine personal trust opinions and global reputations by converging former ones into a single measurement from the whole group, personal aspects are far more complex to express in quantifiable ways than in multi-agent systems. Because users themselves bear some responsibility in contributing to reputation management, most of them are either technically sufficient or would rather spend dedicated time doing so. A negative example is Orkut,* which is used to allow people to express feelings about fellow friends through a rating system known as "karma points," but was finally abandoned due to lack of popularity. Therefore, most live reputation systems of OSNs are based on simple models. Testimonials are still the most popular method for providing member reputation, for example, the number of successful transactions on Ebay and customers' reviews on Amazon.

3. *Search component:* The search refering social graph will leverage performance [32]. Based on the development of Web search techniques [33], searching information in OSNs obeys small world search strategies [34]: using local information about their close contacts [35], for example, when users search jobs on LinkedIn, the results will be ordered by degrees that represent social distances to HRs.

Milgram [34] also explicits a greedy algorithm for small world searching: i will select its neighbor j who is closer to the target t in terms of social distance Y; that is, $Y_{j,t}$ is minimized over all j in i is Web of connections. The problem is developed by Kleinberg [36]. His often cited paper proposed a decentralized search algorithm to solve searching in small worlds with partial information. Kleinberg [37] also provides a theoretical foundation by proving efficient searchability in social networks: a simple algorithm that combines information of network connections, therefore social identities can succeed in efficient searching. **Q3**

Another principal character of searching in OSNs is based on tagging. The content is semantically annotated for better understanding and searchability, for example, tag cloud has recently been utilized by most OSNs. Our paper will discuss this issue in the semantic web part of Section 18.5.

18.3.3 Social Presence Services

Most Social Presence Services in OSNs are created by users to enhance their impression, for example, adding multimedia content, or modifying the look and feel of their page.

* www.orkut.com.

Some OSNs, such as Facebook and Orkut, allow users to add modules or applications that publish various content and interact with others. All that data describe the nature of individuals' presence in the OSNs. The data also enhances their sense of self-worth and stimulates actors to maintain passive interaction within them.

Apart from interactions with their friends on SNs, actors of OSNs tend to present them by current state and activities. Therefore, OSNs offer different mechanisms to support social presence, such as custom messages, online status, the listening of music, watched movies, and so on. By assembling all that data, other actors can form an overall impression of his/her presence in the OSNs and even some clue on what he/she is like in the real world.

There are two types of mechanisms for Social Presence Services. The first type provides social presence by user-generated content, for example, Bob posts a microblogging entry saying "I am feeling good." The second type, context sensitivity, automatically obtains social presences from individuals' real-world context, such as location and time. A case in point is Google Latitude*, which enables a user to allow connected people to track their location. The main difference between context sensitivity and other interactions is that the attention paradigm is reduced to shorter time periods [13]: Compared with interactions such as a direct chat or e-mail, context sensitivity aware friends with no cost of time. In the following paragraphs, we will give detailed explanations and examples of both types.

Q4

1. *User-generated content:* One principal character of OSNs is that most content is user-generated. User-generated Content (UGC) refers to online media content that is produced by end-users. For example, Wikipedia, a web encyclopedia which has 14 million articles†, is written and edited by its users, who can be anyone with access to the site. Admittedly, UGC services are not necessarily SNSs. Considering that OSNs rely on content created by its users who update profiles, communicate with friends, and participate in communities. Integrated UGC services could boost the usage of OSNs through improved user engagement. On the other hand, utilizing OSNs to support online collaboration could improve the quality of UGC [38]. Although UGCs can be created with little or no restriction, monitoring and administration are also necessary to avoid offensive content, copyright issues, and so on.
   Microblogging is a brief text blogging that allows users to send blogs in limited length. It is inspired by cell phone SMS. It can also update multimedia, called micromedia, such as photos or audio. Its most distinguished character is a variety of means of submitting: web, text messaging, instant messaging, email, and so on. Another character is brief: a single sentence or a very short video. By congregating short entities, a logs of the daily events are presented.
   Social news is a team that refers to web services, in which users can submit and vote on news stories. Compared with formal news, which is published by a media agency, social news is collected and edited by end-users. Two of the most popular

* www.google.com/latitude/
† en.wikipedia.org/wiki/Wikipedia.

social news websites are Slashdot and Digg. Today, even media giants such as, CNN start to adapt to social sews.*

Social bookmarking is a service used to share, organize, search, and manage bookmarks of Web resources, typically a Web page. Social bookmark services also encourage users to add, modify, and remove annotations on Web pages. Annotated Web pages will also be visible to other users who share similar annotations which indicate their similar interests. Clouds of tags is a team used to describe clusters of tags or bookmarks provided to users dynamically according to their interests. Social bookmarking services also provide feeds for their lists of bookmarks to allow subscribers to become aware of new bookmarks [39]. For example, users of Google Reader† can either share and tag interested Web pages or subscribe to friends' feeds. Some extra features, such as ratings, may be added to social bookmarking services, for example, research in semantic Web [40,41] propose social bookmarking systems, which are embedded by more semantic means. By referring ontology knowledges in the real world, such as hierarchical relationships.

Wiki is a collaborative encyclopedia service that allows any user to contribute content by creating and editing Web pages. Wikis may serve different purposes, such as learning, collaboration, and knowledge-sharing. Wikis can facilitate social processes [42] in the sense of rewarding contribution [43]. Wiki softwares are also used in corporate intranets, mostly as knowledge-management systems, but in terms of online services, Wikipedia is the most famous and among one of the most typical examples of Web 2.0 services. Wikipedia reached three million english articles in August 2009 and enjoyed the title of being "the largest encyclopedia in the world."

2. *Context sensitivity:* Context sensitivity is a set of SNSs that provide services considering sensed information of users. It became popular greatly because of the rise in using mobile devices for social networking. Contrary to manual implications of personal and contextual information, mobile devices can automatically obtains users' information by various sensors, such as GPS. With the mention of coupling gathered data automatically, such as time coupled with location, or personal information coupled with location, a new service could be created. For example, location sensitivity can provide localized information according to users' location. Another example is personal information coupled with health information [44].

18.3.4 Social Interaction Services

Unlike social presence services, which are content-focused, social interaction services exist mainly for helping other people and for carrying activities to increase the senses of affiliation, belonging, and community.

* CNN iReport, www.ireport.com.
† reader.google.com.

1. *Multiplayer Online Games:* MOGs, such as Second Life and World of Warcraft, enact social networking processes by providing an avatar for game players. They can communicate and live along-side other players. Some OSNs also provide light-weight MOGs as part of their services.
2. *Reviews and Opinions:* Online feedback services mainly provide two functions. First, they publish reviews and opinions to beware public users, which affect customers' decision to buy a certain product. Second, they collect reviews and options from end-users by providing them facilities for writing their own personal views, such as satisfactions or dissatisfactions. Combining social networking with ROS will allow customers to compare products based on reviews from their own connections.
3. *Finance:* Finance services, such as micropayment, provide a system for members to pay for tangible or virtual goods. It is capable of handling arbitrarily small amounts of money. Micropayments have to be suitable for the sale of non-tangible goods over the Internet. This imposes requirements on speed and cost of processing the payments: delivery occurs nearly instantaneously on the Internet, and often, in arbitrarily small pieces. On the other hand, OSNs can easily obtain reputation references to support micropayment.
4. *Groups:* A group is a loosely coupled system of mutually interacting interdependent members. But, a group is more than just the collection of members. Groups can be defined by psychological and temporal boundaries, interact with each other, and task and goals [45]. Most online communities grow slowly at first, due in part to the fact that the strength of motivation for contributing is usually proportional to the size of the community. As the size of the potential audience increases, so does the attraction of writing and contributing. This, coupled with the fact that organizational culture does not change overnight, means creators can expect slow progress at first with a new virtual community. As more people begin to participate, however, the aforementioned motivations will increase, creating a virtuous cycle in which more participation begets more participation.

Q5

18.4 Online Social Network Platforms

This section introduces and compares several popular OSNPs. Despite there being many main players on our candidate list, we select the platform that is distinguished either in its features or ability to represent its own class. One exception may be Facebook, with more than 250 million active users and nearly half of them logging on more than once per day. Despite Facebook's popularity, the potential of OSNs is still far beyond reach and greatly relies on its divided types and separated target markets. As we will discuss later, such diversity promises us that OSNs play or will play in almost every aspect of our lives.

Generally, OSNPs attempt to provide various SNSs in one platform to achieve diverse requirements. Namely, we categorize them as general purpose OSNPs, such as Myspace and Facebook, and other OSNPs that serve narrower target markets called Niche Communities [5], such as LinkedIn. Their different business models end up becoming

divided structures and technical specifications. The rise of OSNs shifts the online community from Websites dedicated to interests, into Websites organized around people. According to social sciences, the Web became an ego-network [46] composed of a person (social actor), friends and family members (social ties), and other people a person knows without personal emotion (social alter). Based on different social network components, the second classification perspective is based on various focuses. We propose this via classified OSNPs into three categories:

Social media is a type of OSNP mainly congregated around SNSs that aim for mass communication, like other news media. Notably, traditional news media, such as CNN and BBC, relay through broadcasting (one to many), whereas social media are more decentralized and rely on dialogues between users (many to many).

Social interaction includes various types of combinations between friends or well-known people. It forms a clique of people who use OSNs as an extension of their social interaction platform, such as LinkedIn or alumni sites. In entertainment, the virtual worlds can be classified into this kind, such as Second Life and the Sims Online.

Social networking: OSNs focus on developing social ties or maintaining the existing ones. In other words, making friends out of strangers. Here, *Networking* emphasizes relationship initiation, unlike social interaction, which is normally between close friends. Social networking focuses on relationships built with strangers [5]. The most common types are dating sites, such as eHarmony and a business relationship site such as LinkedIn. **Q6**

Admittedly, the above classifications are somehow fuzzy and some platforms may be hybrids of two or more types. One example is LinkedIn, which is for both social interaction and social networking: although its main function is seeking a job—a social interaction platform. LinkedIn also plays another important role: networking HR and former colleagues. Next, we will present several live OSNPs and further explain our points on common or unique characters of OSNPs.

LinkedIn is a business-oriented social networking site. Founded in December 2002 and launched in May 2003, it is mainly used for professional networking and has become a powerful business service. As of October 2009, LinkedIn had more than 50 million registered users, spanning more than 200 countries and territories worldwide. LinkedIn controls what a viewer may see based on whether the viewer has a paid account. LinkedIn allows users to opt out of displaying their network. Compared with other OSNPs, LinkedIn's business model is unique, in which it charges the user for accessing personal information.

Flickr [47] is an image and video-hosting website, Web services suite, and an OSN. Flickr provides both private and public image storage. A user uploading an image can set privacy controls that determine who can view the image. A photo can be flagged as either public or private. Private images are visible by default only to the uploader, but they can also be marked as viewable by friends and/or family. Privacy settings can also be decided by adding photographs from a user's photo stream toa "group pool." If a group is private, all the members of that group can see the photo. If a group is public, the photo becomes public as well. Flickr also provides a "contact list" that can be used to control image access for a specific set of users in a way similar to social tier tools of other OSNs.

Facebook is the world's largest social network, with over 350 million active users and half of them visiting the site once per day*. It basically provides a platform to share a common interest, idea, task, or goal within its users, where they are able to develop or maintain personal relationships. Moreover, it also provides applications of various services, such as social bookmarking and instant messaging. Like other social networks, the site allows its users to create a profile page and forge online links with friends and acquaintances. Facebook launched its API in 2007, providing a framework for software developers to create applications that interact with core Facebook features. But, its API placed several restrictions on having complete access to an individual's social graph.

Ning was launched in October 2005. It is an OSNP for people who want to create their own social networks. Ning competes with social sites like MySpace and Facebook by appealing to people who want to create their own social networks around specific interests with their own visual design, choice of features, and member data [5]. The unique feature of Ning is that anyone can create their own social network for a particular topic or need, catering to specific membership bases. Ning has both free and paid options to fully eliminate advertisements. When someone creates a social network on Ning, it is free by default and runs ads that Ning controls. If the person creating the social network chooses, they can pay to control the ads (or lack thereof) in exchange for a monthly fee. A few other premium services, such as extra storage and bandwidth and non-Ning URLs, are also available for additional monthly fees. However, Ning does allow developers to have some source level control of their social networks, enabling them to change features and underlying logic.

Realtravel tries to solve the problem: how to extract information from data. In a collective knowledge system, the aggregate content must be more useful: create aggregate values by integrating user contributions of unstructured content with structured data. RealTravel attracts people to write about their travels to share stories and photos with semantic annotation. Travel researchers enjoy the benefit of all experiences relevant to their target destinations.

18.5 Research Topics and Challenges

18.5.1 Key Technologies

Broadly speaking, there are two types of concerns in OSN research: data access issues and data publication-related issues. In this section, we will discuss the following types of research topics as well as challenges:

- Distributed architecture;
- Fragmented user identity;
- Contextual information associated users and possible abuse;
- Identity and trust;
- Policies within network and Web of trust (*dilemma: usability vs. privacy);
- Deeper adaptive user experiences.

* http://www.facebook.com/press/info.php?statistics.

We will also cover some challenges of different social aspects: bridging online and offline social networks, positive interactions, and so on.

Distributed architecture: One key question pertaining to architecture of OSNs is whether a decentralized architecture is sustainable, profitable, and usable, and consequently, what do we stand to lose if we adopt a decentralized architecture. Considering fragmentation of Web capabilities, how to avoid overhead in processing information in OSNs? What is the minimum set of new functionalities that the future web should incorporate?

In contrast to the increasingly sophisticated capabilities of services, the fundamental architecture of the Web has not changed much over the past 10 years. Existing social networks usually employ a "hub and spoke" model, where the website is the hub of all activity within the network, and where there is a *client* and a *server*. Since all traffic must pass through the hub, that site may become a bottleneck. Furthermore, each transaction must pass up one spoke to the hub, and then down another spoke, when the people interacting may be much closer to each other; in network terms, than either is to the hub site.

Services and applications in OSNs have become quite sophisticated in the features they provide. There is an opportunity to create an architecture that distributes the load. Such an architecture would require better interoperability between OSNs, more-so than what we have available, and should remove any dependence on an "always-on" network connection.

However, the hurdles for distributed OSNs are great, some being fundamental, such as incompatible assemblies, different data access APIs, and the entity data model. Despite difficulties, some prototypes of distributed OSNs have developed. The appleseed project is an open source OSN framework which is based on a distributed model. For instance, a profile on one Appleseed Website could "friend" a profile on another Appleseed Website, and the two profiles could interact with each other. **Q7**

Privacy and trust: For OSNs, identities and links are more important than content. The privacy issue on OSN can be classified into three types: identity disclosure, link disclosure, and content disclosure.

A report* finds that over half (52%) of social network users post risky information online. For example, the report states that 73% of adult Facebook users shared content only with friends, but only 42% of users state that they customized their privacy settings.

When using Web-based social networks to refer trust values, most of the information that sociology considers important is not available (e.g., we do not know the history between people, the user's own background, and how likely they are to trust, in general, the familial/business/friend relationship between users). Thus, we must understand trust only from the available information. Privacy is also implicated in users' ability to control impressions and manage social contexts. Boyd (in press-a) asserted that Facebook's introduction of the "News Feed" feature disrupted students' sense of control, even though data exposed through the feed was previously accessible. Some research argued that the privacy options offered by OSNs do not provide users with the flexibility they need to handle conflicts with Friends who have different conceptions of privacy;

* State of the Net 2010-Consumer Reports.

they suggest a framework for privacy in SNSs that they believe would help resolve these conflicts.

Identity and profile: Most research in this subject focuses on enhancing user security without compromising usability. The question left is, how to mirror OSNs with users' real identities?

To answer this question, we need to dive into reality first—identity mapping and wall barriers. Since each user has many registrations or accounts, the attention is dispersed. Identity in different OSNs exists as separate, isolated islands of discourse, unable to exchange meaningful information, leverage their accumulated knowledge, or connect with other communities that share their concerns. As the user takes a more active role in the production of content, and even services, and becomes a "prosumer," this situation leads to a somehow chaotic scenario where the same user is present in an uncountable number of different platforms, taking the best-of-breed for any aspect of social interaction or simply following or joining their friends. This situation creates an increasingly inconvenient and uncomfortable situation where users not only own different accounts, each one with a specific set of credentials, but also deal with an increasing amount of personal information scattered throughout several sites, each with different data usage policies and privacy protection conditions. Finally, how can we allow users who may want to deliberately fragment their online identity to do so?

There are also several independent initiatives focusing on how to break the wall by providing persistent identity. They first appeared as liberty alliances, such as Microsoft's .Net identity system named .Net Passport originally, and changed into Live Passport. Microsoft had accumulated various services, such as the Hotmail and MSN Spaces. But such effort faced significant resistance from other companies and users. There is great concern that online identity might become the property of a single corporation. Such centralized control would be devastating. As a result, vender-neutral identity services emerged, such as OAuth and OpenID. They both provide an open protocol to allow secured API authorization in a simple and standard method. Similarly, OAuth allows using anonymous tokens instead of usernames and passwords as identity. The granularity of permission can be either site level or application level, even a defined duration. OAuth can also grant a third party site access to their information stored with another service provider, without sharing their access permissions or the full extent of their data.

Q8

Structured data: According to the collective intelligence theory from Doug Engelbar: The grand challenge is to boost the collective IQ of the organizations and of the society. To achieve this, the information on the web has to be structured. Semantic Web is used to define information and services on the Web, making them possible for the Web to "understand" and satisfy the requests of people and machines that use the Web content. According to Tim Berner-Lee, "The Semantic Web is not a separate Web, but an extension of the current one, in which information is given a well-defined meaning, better enabling computers and people to work in cooperation."

One attempt is FOAF + SSL. It is a machine-readable ontology describing persons, their activities, and their relations with other people and objects. Anyone can use FOAF to describe him or herself. FOAF allows groups of people to describe social networks without the need for centralized databases. FOAF is a descriptive vocabulary team

expressed using the Resource Description Framework (RDF) and the Web Ontology Language (OWL). Computers may use these FOAF profiles to find, for example, all people living in Europe, or to list all people both you and a friend of yours know. This is accomplished by defining relationships between people. Each profile has a unique identifier (such as the person's e-mail addresses, a Jabber ID, or an URI of the homepage or Weblog of the person) which is used when defining these relationships.

Other efforts include microformats, such as XFN and hCard. It is a Web-based approach to semantic markup that seeks to re-use the existing XHTML and HTML tags to convey metadata and other attributes. This approach allows information intended for end-users, such as contact information, geographic coordinates, calendar events, and the like to also be automatically processed by the software. Unlike the formal semantic Web, which is more complex, the microfomat is light-weight and easy to implement in even today's Web markup languages, for example, HTML5 adapts to several microformats.

Mobile social networking: Mobile social networking is a concept combining mobile communication and social networking. To illustrate the scale of mobile social networking, the number of unique visitors to the Facebook mobile site increased fivefold from 5 million per month in January 2008 to 25 million per month in February 2009. The latter figure represents 18% of Facebook's 120 million users (February 2009), a proportion that has gradually increased over time, and it will continue to do so in the coming years. Social networks with an established presence on the fixed line Internet are clearly benefiting from extending their services over mobile channels.

One obvious advantage of mobile social networking is context sensitivity, which means, in terms of places, time and people makes services more information-sensitive. Mobile devices can collect more personal information than normal PCs, such as locations and contacts. By adding various sensors into mobile devices, new types of applications can go beyond the existing domains. Location-based services (LBSs) are among the most popular ones. An LBS is an information and entertainment service, accessible using mobile devices through the mobile networking, which utilizes the ability to make use of the geographical position of the mobile device. It can be used in a variety of contexts, **Q9** such as health, work, personal life, and so on. LBS services include services of identifying a location of a person or object, such as discovering the nearest ATM or the whereabouts of a friend or employee. LBS services include parcel-tracking and vehicle-tracking services. LBS can include mobile commerce when taking the form of coupons or advertising directed at customers, based on their current location. They include personalized weather services and even location-based games.

Accessibility and user experience: How to deepen and adapt user experiences are also important practices of making OSNs useful. Human factors can be explicated in two levels: the general user experience and especially UI guidelines for accessibility. Accessibility means how the information in the OSNs can be correctly built and maintained, so all of these users can be accommodated while not impacting the usability of the site for non-disabled users. User experience covers a wider context of how to capture and better support social activities. For example, an OSN focus or hobby focus can be treated as a self-organized system, in which global patterns emerge from local actions and structured subsequent local actions.

The challenge in this research is in the mapping of quantitative measurements of interactions based on network traffic to qualitative analyses of social relationships. It is easy to know what people are doing in the network, but it is harder to know why. Most research is empirical and their fundamental theories are beyond the scope of our survey.

18.6 Related Concepts

As the increase in the popularity of OSNs constantly rises, academic research is emerging from diverse disciplinary and methodological information systems that can take advantage of the users' social and personal data, address a range of topics, and build on a large body of social network research. Broadly speaking, research on social networking can be divided into: how to effectively and positively communicate in OSNs,

After Milgram's study revealed small world property in social networks, research also found that social networking shared common characters, such as weak ties, power-law, and fuzzy boundaries. As a sociology concept, social networking is a social structure composted by individuals and relationships within them. An OSN or virtual community, however, is an internet-based community and information system of social networking. The idea of social networking is both old and new. Although it is a common phenomenon existing in every human interaction, when we talk about Social Web, in this paper, we intend to focus on online social networking sites, which are also called "online social networks" or "virtual communities". But the theories may build on each other.

18.6.1 Social Web

The concept of Social Web that research expected, is the web of people. It shares the same features of real social relationships, such as six degrees of separation phenomenon, scale-free, and so on. Unlike OSNs, which maintain weak ties, strength, and latent tie, social web has its limitations, for example, one factor is Dunbar's number, which points out the "theoretical cognitive limit to the number of people with whom one can maintain stable social relationships," is generally accepted to be about 150.

The "Six Degrees of Separation" phenomenon was first investigated by Stanley Milgram [48] in 1960, where he addressed letters to a particular stockbroker in New York and gave them to people, randomly picked at locations in the United States, far away from that of the final receiver. The condition for passing the letter, so that it reaches the addressee, was that one could post it only to people they knew personally by first name. Eventually, most of the letters reached the destination and the average number of hops was six. Since then, there have been various studies demonstrating how this effect may help people conduct their everyday lives.

This effect, also known as *Small Worlds* or *Scale*-Free *Networks*, has been revisited with analytical techniques starting with the seminal work of Barabasi [4,8,28]. Barabasi studied many natural and man-made networks and found that they all exhibit degrees of clustering with hub and spoke topologies and remote links betweens clusters. These real networks are fundamentally defined by a few highly connected nodes, but even a very small number of remote links (weak ties) are sufficient to dramatically decrease the average separation between nodes. Analytically, Barabasi measured this clustering effect with

power-law distributions, showing varying power law exponents for networks, such as movies (by their actors especially Kevin Bacon), members of an audience (through auditory cues), social systems (family ties, school ties, friendships, etc.), biological organisms (biochemical signals), the brain (neural interconnections), and especially, the Internet. In fact, there have been a series of studies of the structure and topology of Internet-based networks best summarized in [25], including the web, email, instant messaging, virus/worm propagation, and P2P networks. Before this work, identifying and quantifying the scale-free nature of the Internet, every new algorithm proposed by researchers for improving network performance was typically tested on random networks generated by consensus tools (such as the Waxman Network Topology Generator*), which in retrospect, resulted in incorrect solutions, which should now be re-examined.

One set of concepts related to OSNs are strong ties, weak ties, and latent ties [49]. Some research explains why relationships in OSNs are weak ties. The positive effect of it is: "communities of interest are defined by their worldviews, and whenever a community of interest rigorously exposes its worldview in a fashion that permits its knowledge to be federated with the worldviews and knowledge of other communities, the whole human family is enriched"—Steven Newcomb. The research designed for positive social change also found that OSNs may differ in purpose, but their architectures and interactive patterns share a lot in common [50]. Such topics have been extensively studied in theoretical works, such as complex network theory [51].

18.6.2 ERP

Finally, we will explicit some systems, which are not OSNs, but share some characteristics. ERP systems, such as customer relationship management system and human resources management system, are basically role-driven. An user gets a role and a responsibility for the quality of the data in the process. There is no consistency in relationships. They are initiated by an individual sending a request for participation in a narrowly defined project, and would be forwarded based on expressed affinities and the recommendations of trusted third parties. The resulting ad hoc community would dissolve with the completion of the stated objective. **Q10**

Another similar business system is groupware, which is a software systems for collaborating within a group, such as email, calendaring, text chat, and wiki. Despite the notion of collaborative work systems, which are conceived as any form of human organization that emerges any time collaboration takes place, whether it is formal or informal, intentional or unintentional. In normal terms, it is business software and not public accessible OSNs.

18.7 Conclusion

The work described above is an ongoing dialogue for both practitioners and researchers. New social network services emerge every day. The platforms we analyzed adjust themselves continuously. Methodologically, we can only make causal claims, is limited by a

* http://www.math.uu.se/research/telecom/software/stgraphs.html\.

snapshot of the development of OSNs. Our work surveys the Web services combining social networks and information system, and leverages the advantages of each type of system. We noticed that most current ONSs implemented only very simple models of social networking and cannot mirror the richness of real-world complexity. On the other hand, due to either technology restrictions or business concerns, the big players in the market cannot, or would not, open their platform to achieve the full potential of OSNs. We hope our survey can advocate a future research agenda to melt the gap.

References

1. Cross, R., Parker, A., Borgatti, S. P. A bird's-eye view: Using social network analysis to improve knowledge creation and sharing. *Knowledge Direct* 2000; 2(1): 48–61.
2. Loscalzo, S., Yu, L. Social network analysis: Tasks and tools, 2008; 151–159.
3. Scott, J. P. *Social Network Analysis: A Handbook*, Sage Publications Ltd, May 2000.
4. Granovetter, M. S. The strength of weak ties. *Am J Sociol* 1973; *78*(6): 1360–1380.
5. Boyd, D., Ellison, N. Social network sites: Definition, history, and scholarship. *J Comput-Mediat Commun* 2008; *13*(1): 210–230.
6. Smith, M., Kollock, P. *Communities in Cyberspace*, 1 ed., Routledge, February 1999.
7. Ziegler, C-N., Lausen, G. Propagation models for trust and distrust in social networks. *Information Syst Frontiers* 2005; 7(4–5): 337–358.
8. Haythornthwaite, C. Strong, weak, and latent ties and the impact of new media. *The Information Soc* 2002; 18: 385–401.
9. Biocca, F., Harms, C., Burgoon, J. K. Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence: Teleoperators Virtual Environ* 2003; *12*(5): 456–480.
10. Java, A., Song, X., Finin, T., Tseng, B. Why we twitter: understanding microblogging usage and communities, in WebKDD/SNAKDD'07: *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis*, New York, NY, USA, 2007, ACM, pp. 56–65.
11. Cacioppo, J. T., William, P. *Loneliness: Human Nature and the Need for Social Connection*. Norton Press, New York, 2008.
12. Oreilly, T. What is web 2.0: Design patterns and business models for the next generation of software. Social Science Research Network Working Paper Series, August 2007.
13. Bentley, F., Metcalf, C. J. The use of mobile social presence. *IEEE Perv Comput* 2009; 8(4): 35–41.
14. Short, J., Williams, E., Christie, B. *The Social Psychology of Telecommunications*. John Wiley and Sons Ltd, 1976.
15. Chung, D., Debuys, B., Nam, C. Influence of avatar creation on attitude, empathy, presence, and para-social interaction, 2007; 711–720.
16. Stets, J. E., Burke, P. J. Identity theory and social identity theory. *Social Psychol Q* 2000; 63(3): 224–237.
17. Boyd, D. M. Friendster and publicly articulated social networking, in *CHI '04: CHI '04 Extended Abstracts on Human factors in Computing Systems*, New York, NY, USA, 2004, ACM Press, pp. 1279–1282.

18. Baatarjav, E-A., Dantu, R., Phithakkitnukoon, S. *Privacy Management for Facebook*, 2008, pp. 273–286. **Q13**
19. Backstrom, L., Dwork, C., Kleinberg, J. Wherefore art thour3579x?: anonymized social networks, hidden patterns, and structural steganography, in *WWW '07: Proceedings of the 16th International Conference on World Wide Web*, New York, NY, USA, 2007, ACM, pp. 181–190.
20. Liu, K., Terzi, E. Towards identity anonymization on graphs, in *SIGMOD '08: Proceedings of the 2008 ACM SIGMOD Internationa Conference on Management of Data*, New York, NY, USA, 2008, ACM, pp. 93–106.
21. McCown, F., Nelson, M. L. What happens when facebook is gone? in *JCDL '09: Proceedings of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries*, New York, NY, USA, June 2009, ACM, pp. 251–254.
22. Jordan, K., Hauser, J., Foster S. The augmented social network.
23. Brzozowski, M. J., Hogg, T., Szabo, G. Friends and foes: ideological social networking, in *CHI '08: Proceeding of the Twentysixth Annual SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2008, ACM, pp. 817–820.
24. Levien, R., Attack-resistant trust metrics, 2009, 121–132.
25. O'Donovan, J. Capturing trust in social web applications, 2009, 213–257.
26. Pitsilis, G., Marshall, L. Modeling trust for recommender systems using similarity metrics, 2008, 103–118.
27. Ruohomaa, S., Kutvonen, L. *Trust Management Survey*, vol. 3477, 2005, pp. 77–92.
28. Guha, R., Kumar, R., Raghavan, P., Tomkins, A. Propagation of trust and distrust, in *WWW '04: Proceedings of the 13th International Conference on World Wide Web*, New York, NY, USA, 2004, ACM, pp. 403–412.
29. Kamvar, S. D., Schlosser, M. T., Garcia-Molina, H. The eigentrust algorithm for reputation management in p2p networks, in *Proceedings of the 12th International Conference on World Wide Web (WWW '03)* (New York, NY, USA, 2003), ACM, pp. 640–651.
30. Page, L., Brin, S., Motwani, R., Winograd, T. The pagerank citation ranking: Bringing order to the web, Technical report, Stanford Digital Library Technologies Project, 1998.
31. Massa, P., Avesani, P. Controversial users demand local trust metrics: an experimental study on epinions.com community, 2005.
32. Chi, E. H. Information seeking can be social. *Computer* 2009; 42(3): 42–46.
33. Chakrabarti, S. *Mining the Web: Discovering Knowledge from Hypertext Data*, 1st ed., Data Management Systems, Morgan Kaufmann, August 2002.
34. Travers, J., Milgram, S. An experimental study of the small world problem. *Sociometry* 1969; 32(4): 425–443.
35. Adamic, L. A., Adar, E. How to search a social network. 2004.
36 Kleinberg, J. The small-world phenomenon: An algorithmic perspective. Technical report, Ithaca, NY, USA, 1999.
37. Kleinberg, J. Complex networks and decentralized search algorithms, in *International Congress of Mathematicians (ICM)*, 2006.

38. Hu, M., Lim, E-P., Sun, A., Lauw, H. W., Vuong, B-Q. Measuring article quality in wikipedia: models and evaluation, in *CIKM'07: Proceedings of the 16th ACM Conference on Conference on Information and Knowledge Management* (New York, NY, USA, 2007), ACM, pp. 243–252.
39. Chen, F., Scripps, J., Tan, P-N. Link mining for a social bookmarking website, in *WI-IAT '08: Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology* (Washington, DC, USA, 2008), IEEE Computer Society, pp. 169–175.
40. Kiryakov, A., Popov, B., Terziev, I., Manov, D., Ognyanoff, D. Semantic annotation, indexing, and retrieval. *Web Semantics: Sci, Serv Agents on the World Wide Web* 2004; 2(1): 49–79.
41. Wu, X., Zhang, L., Yu, Y. Exploring social annotations for the semantic web, in *WWW '06: Proceedings of the 15th International Conference on World Wide Web*, New York, NY, USA, 2006, ACM, pp. 417–426.
42. Cress, U., Kimmerle, J. A systemic and cognitive view on collaborative knowledge building with wikis. *Int J Computer-Supported Collab Learning* 2008; 3(2): 105–122.
43. Hoisl, B., Aigner, W., Miksch, S. Social rewarding in wiki systems motivating the community, 2007; 362–371.
44. Morris, M. E. Social networks as health feedback displays. *IEEE Internet Comput* 2005; 9(5): 29–37.
45. Arrow, H., Mcgrath, J. E., Berdahl, J. L. *Small Groups as Complex Systems: Formation, Coordination, Development, and Adaptation*, 1st ed., Sage Publications, Inc, March 2000.
46. Lin, N. *Social Capital : A Theory of Social Structure and Action*, Cambridge University Press, 2001.
47. Mislove, A., Koppula, H. S., Gummadi, K. P., Druschel, P., Bhattacharjee, B. Growth of the flickr social network, in *WOSP '08: Proceedings of the first workshop on Online social networks,* New York, NY, USA, 2008, ACM, pp. 25–30.
48. Milgram, S. The small world problem. *Psychol Today* 1967; 2: 60–67.
49. Viswanath, B., Mislove, A., Cha, M., Gummadi, K. P. On the evolution of user interaction in facebook, in *WOSN '09: Proceedings of the 2nd ACM workshop on Online social networks*, New York, NY, USA, 2009, ACM, pp. 37–42.
50. Kumar, R., Novak, J., Tomkins, A. Structure and evolution of online social networks, in *KDD '06: Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA, 2006, ACM, pp. 611–617.
51. Mislove, A., Marcon, M., Gummadi, K. P., Druschel, P., Bhattacharjee, B. Measurement and analysis of online social networks, in *IMC '07: Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement*, New York, NY, USA, 2007, ACM, pp. 29–42.