Part I

Introduction

Cambridge University Press & Assessment
978-0-521-19727-4 — Behavior Dynamics in Media-Sharing Social Networks
H. Vicky Zhao , W. Sabrina Lin , K. J. Ray Liu
Excerpt
More Information

1 Introduction to media-sharing social networks

With recent advances in communications, networking, and computer technologies, we have witnessed the emergence of large-scale user-centered web 2.0 applications that facilitate interactive information sharing and user collaboration via Internet – for example, blogs; wikis; media-sharing websites such as Napster, Flickr, and YouTube; social networking services such as Facebook, LinkedIn, and Twitter; and many others. Different from traditional web applications that allow only *passive* information viewing, these web 2.0 sites offer a platform for users to *actively* participate in and contribute to the content/service provided. The resulting trend toward social learning and networking creates a technological revolution for industries, and brings new experience to users.

The emergence of these websites has significant social impact and has profoundly changed our daily life. Increasingly, people use the Internet as a social medium to interact with one another and expand their social circles, to share information and experiences, and to organize communities and activities. For example, YouTube is a popular video-sharing website on which users upload, share, and view a wide variety of user-generated video content. It targets ordinary people who have Internet access but who may not have a technical background on computers and networking, and enables them to upload short video clips that are viewable to the worldwide audience within a few minutes. Its simplicity of use and the large variety of content offered on the website attract more than one billion views per day, according to a blog by Chad Hurley (cofounder of YouTube) on October 9, 2009, and make video sharing an important part of the new Internet culture.

According to the Alexa Global traffic ranking, among the 20 hottest websites, many of them are social networking and media-sharing websites – for example, Facebook, MySpace, Twitter, and YouTube. The increasing popularity of these interactive and user-centered websites also means new business opportunities. In July 2009, eMarketer projected that even with the current world economy hurdle, the amount of money US marketers spend on online social network advertising will reach $1.3 billion in 2010, a 13.2 percent increase compared with 2009 [1]. In a later report in December 2009, eMarketer predicted that in 2010, worldwide online advertising spending on Facebook would reach $605 million, corresponding to a 39 percent increase compared with 2009 [2]. In particular, it was predicted that the non-U.S. advertising spending on Facebook would increase by 65 percent in 2010 [2].

With recent advances in wireless communication technologies, mobile social networks have become increasingly popular. Internet-based social networks such as MySpace and Facebook have turned mobile; they enable mobile phone users to access their websites, upload mobile photos and videos to their profiles, and so forth. In addition, new mobile social networks that are designed specifically for mobile applications – for example, Foursquare, Loopt, and Gowalla, which enable users to explore and discover their local vicinity – have also emerged. A report released by Informa on March 25, 2010 forecast that US mobile social networking ad revenue would rise by 50 percent to $421 million in 2010, and would continue its robust growth into 2013 with a breakthrough of the $1 billion revenue mark [3]. Other potential applications for integrated sensor and social networks include traffic monitoring, human movement and behavior analysis, collaborative rehabilitation for seniors, and many others [4,5].

However, this emerging trend of social learning and networking also poses new challenges and raises critical issues that need to be addressed for further proliferation and development of such social networks. Listed below are just a few of them.

- With the resulting avalanche of information that users create, share, and distribute over networks, it is crucial to effectively manage these data and to support accurate and fast searching of information. This is particularly challenging for media objects (audio, image, and video), as the same audio or video clip (or its portions) may be processed in many different ways and appear in a variety of different contexts and formats [6]. As an example, a recent study of traffic flow in FastTrack, one of the largest peer-to-peer (P2P) file-sharing networks, showed that there were 26,715 different versions and 637,381 different copies of the song "Naughty Girl" on FastTrack. Among them, 62 percent of the versions and 73 percent of the copies were "polluted" – that is, either they were nondecodable or their time durations were significantly longer or shorter than the official CD release [7].
- Social networks facilitate easy information sharing among users, and the same easy access of such networks enables *anyone* to view the shared content [8]. Information sharing at such an unprecedented scale in social networks may post serious security and privacy concerns. For example, in medical and scientific research, collecting human behavior and health information requires very strict scrutiny, whereas social networks make it possible to collect such information much more easily without contacting the subjects [9]. Using and republishing such public information in research without informed consent may be considered as an invasion of privacy [9].
- Social networks may be misused and manipulated by people for defamation, profit, and many other purposes [10]. For example, researchers from Harvard University recently discovered that scammers created sophisticated programs to mimic legitimate YouTube traffic and to provide automated feedback for videos and other content they wished to promote [11]. A recent study by researchers at the University of California at Berkeley found that some eBay users artificially boost their reputations by buying and selling feedbacks, so they can seek higher prices on items that they sell [12].

From these examples, it can be seen that users play an active and important role in social networks, and the way in which they behave and use the huge amount of information available on social networks has a significant impact on system performance. These new challenges call for novel solutions to model user interactions and study the impact of human behavior on social networks, to analyze how users learn from each other as well as from past experiences, and to understand people's cognitive and social abilities. These solutions will facilitate the design of future societies and networks with enhanced performance, security, privacy, availability, and manageability. This is an interdisciplinary research area, covering signal processing, social signal processing, information science, sociology, psychology, and economics, in which signal and information processing plays a critical role. The advanced signal and information processing technologies will enable us to better characterize, understand, and ultimately influence human behaviors as desired.

This book focuses on an important class of social networks, media-sharing networks, in which users form a dynamically changing infrastructure to upload, exchange, distribute, and share images, videos, audio, games, and other media. Famous examples include YouTube, Napster, and Flickr. Also, many P2P file sharing systems – for example, BitTorrent and KaZaa – have been used to share digital media. Catching the current trend of delivering TV programs over the Internet, we have also seen many successful deployments of P2P live streaming, sometimes called P2PTV, in which video streams (typically TV programs) are delivered in real time on a P2P network. Examples of such P2PTV applications include PPLive, PPStream, SopCast, QQLive from China, Abroadcasting from the United States, and LiveStation from the United Kingdom. They attract millions of viewers, and the aggregated bandwidth consumption may reach hundreds of gigabits per second [13]. In this book, we study user behavior in media-sharing social networks and analyze the impact of human factors on multimedia signal design. We use two different types of media-sharing social networks, multimedia fingerprinting and P2P live streaming, as examples.

Before we move on to the modeling and analysis of user behavior in media-sharing social networks, we first quickly review recent advances in other research areas in media-sharing social networks, including social network analysis and media semantics in social networks.

1.1 Quantitative analysis of social networks

Social networks are defined as "social structures that can be represented as *networks* – as sets of *nodes* (for social system members) and sets of *ties* depicting their interconnections" [14]. The two elements, actors (or nodes) and relations, jointly form a social network. Actors can be individual persons, small groups, formal organizations, or even countries, who are connected to each other via certain relationships, such as friendship, trade, or colleagues. In addition to describing how a set of actors are connected to each other, *social network analysis* describes the underlying patterns of social structure and investigates their impact on individual behavior, as well as analyzing them on the system level [15].

1.1.1 Social network representation, notations, and relationship measures

1.1.1.1 Representation

There are two different methods to represent and analyze a social network, *sociogram* and *sociomatrix* [15]. In a sociogram, graphs and graph theory are used to visually represent and analyze social networks. Here, actors are denoted as points (also called *nodes* or *vertices*) and a relation (tie) is represented using a line (also called an *arc* or *edge*). A sociomatrix uses tabular matrices to depict social networks and to facilitate complex mathematical analysis. Here, an $N \times N$ matrix x is used to represent a social network with N actors, and the element x_{ij} at row i and column j represents the relationship between the ith and jth actors, in which actor i is the initiator and j is the recipient. These two representation methods are equivalent and contain the same information.

Different networks represent different relations, and there are many different types of social networks. If the relation is *nondirected*, or mutual – for example, classmates and colleagues – all lines in the graph representation have no arrowheads, and in the matrix representation, we have a symmetric matrix with $x_{ij} = x_{ji}$. Other types of relations are *directed*, and we have directed graphs with arrowheads in the lines. For example, A trusts B, whereas B may not trust A; therefore, there is one link from A to B but not vice versa.

In addition to directionality, the lines (arcs) in a social network may be binary or measured with different value scales. For example, in simple relations such as classmates or colleagues, there is either a line (presence) or no line (absence) between two nodes, which corresponds to binary networks. For other types of relations, each line not only indicates the *existence* of the relation, but also values the *intensity* of the relation. For example, one actor can rank other actors in the network as "friends," "acquaintances," or "strangers," indicating different levels of relations [15].

To summarize, there are four basic types of social networks: binary nondirected, binary directed, valued nondirected, and valued directed.

1.1.1.2 Notations

Graph theory is often used in social network analysis; thus, we first introduce some basic concepts in graph theory.

Given a social network represented using a graph, a *subgraph* is a subset of nodes and lines, in which all lines in the subgraph must be between pairs in the subgraph. A *walk* is an alternating sequence of incident nodes and lines, which connects the starting and the ending nodes; the *length* of a walk is the number of lines contained in the walk. A *path* is a walk with distinct nodes and lines – that is, every node and every line are visited only once in the walk.

A graph is *connected* if there is a path between every pair of nodes in the graph, and is called *disconnected* otherwise. A node that is not connected to any other nodes is called an *isolate*. A *graph component* is a maximal subgraph that forms a connected graph. In a connected graph, a node is a *cutpoint* if its removal would disconnect the graph, and a line is a *bridge* if its removal would disconnect the graph into two or more components. The notion of cutpoint and bridge is important in network analysis, as networks with cutpoints and bridges are more vulnerable to disruptions than those with many redundant paths to sustain information and resource flows [15].

1.1.1.3 Relationship measures

There are many important relationship measurements in graph theory – for example, nodal degree, geodesic distances, and density – which we will briefly introduce here.

Nodal degree: For a node i in a binary nondirected graph, its *nodal degree* is the total number of lines that are incident with it. With directed graphs, *nodal indegree* and *nodal outdegree* should be distinguished. Nodal indegree is the number of lines received by node i, and nodal outdegree is the number of lines sent by node i. For valued graphs, we can use the mean values of the lines connected to node i to represent its nodal degree. Nodal degree reflects the node's level of involvement in network activities [15], and the mean nodal degree averaged over all nodes shows the aggregate level of activity in the network.

Geodesic distance: The *geodesic distance* between a pair of nodes is the length of the shortest path that connects them. If there is no path between two nodes, then their geodesic distance is infinite or undefined. For directed graphs, the geodesic distance from node i to node j may be different from the geodesic distance from j to i. For example, node i may be able to send a message to j, but not vice versa. Geodesic distance measures the closeness of two nodes and plays an important role in distance-based analysis, such as in clustering analysis [15].

Density of a graph: *Density* measures the extent to which nodes in a graph are connected among themselves. For a binary nondirected graph with N nodes, its density D is the number of lines in the graph (L) divided by the maximum possible lines ($\binom{N}{2}$ when there is a direct link between any pair of nodes in the graph) – that is, $D = L/\binom{N}{2}$. For directed graphs, the denominator is changed to $2\binom{N}{2}$, because for each pair of nodes in a directed graph, there are two possible lines with different directions. For valued graphs, the numerator is replaced by the summation of all lines' values.

1.1.2 Centrality and prestige

An important usage of graph theory is to identify the "most important" actors and/or groups in social networks [15,16]; the concepts of *centrality* and *prestige* quantify an actor (or group)'s prominence (involvement in the network activities) in a network. An individual actor's prominence reflects its visibility to other actors in the networks, and at the group level, it evaluates the divergence of all group members' prominence [15]. The difference between centrality and prestige is whether the direction of lines counts.

In centrality, a prominent actor has many direct links with other actors regardless of the direction, whereas in prestige, a prominent actor receives many incoming lines but does not initiate many outgoing ties.

The most widely used centrality measures are *degree*, *closeness*, and *betweenness*. In this section, we use binary nondirected graphs to illustrate these concepts, and the definitions for directed and valued graphs are available in references [15,16].

Degree centrality: At the individual level, node i's *degree centrality* is defined as its nodal degree normalized by the total number of actors in the graph, and is a real number between 0 and 1. Actors with high degree centralities have more connections with others and higher visibility in the networks.

At the group level, group degree centrality measures the extend to which actors differ in terms of their individual degree centralities, and resembles the standard deviation of the individual degree centralities among group members [15]. When all group members have the same degree centrality, the group degree centrality is zero. In the other extreme case of a star graph, in which one node is connected to all other nodes but there is no connection between any other two nodes, the group degree centrality achieves the maximum possible value.

Closeness centrality: For node i, its *closeness* reflects how quickly it can interact with other actors, such as by communicating directly or via few intermediaries [15]. For node i, its actor closeness centrality index is the inverse of the mean geodesic distances between i and all other nodes in the graph, and it takes the smallest value when node i is directly linked to all others in the network. At the group level, the group closeness centralization index measures the extent to which actors differ in their individual closeness centralities.

Betweenness centrality: The communication between two nonadjacent nodes depend on other actors, especially those who are on the paths between these two. These "other actors" may potentially have some control over the interaction between these two nonadjacent actors, and the *betweenness centrality* concept quantifies how other actors control or mediate the relations between connected nodes [15]. Actor betweenness centrality measures the extent to which an actor lies on the shortest path between pairs of other actors, and the group-level betweenness measures the extend to which this value varies across group members. Detailed definitions and explanations of betweenness centrality can be found in references [15,16].

Prestige is used when it is much more important to specify the initiators and the recipients of relations than just giving mere participation [15]. It measures the extent to which an actor "receives" relations sent by others, and emphasizes inequality in control over information and/or resources. For node i in a directed graph, its actor degree prestige is its normalized indegree, and it takes a larger value when node i is more prestigious. A detailed discussion can be found in reference [16].

1.1.3 Cohesive subgroups

Another important task in social network analysis is to identify *cohesive subgroups* of actors who are connected via many direct, reciprocated choice relations, and who share information, achieve homogeneity of thoughts and behavior, and act collectively [15,16].

In graph theory, *clique* is an important concept for analyzing group structures and for understanding how cohesion benefits group members as well as restricts the range of social contacts [17,18]. A clique is "a maximal complete subgraph of three or more nodes, all of which are directly connected to one another, with no other node in the network having direct ties to every member of the clique" [15]. Thus, every pair of nodes in a clique is connected by a direct link, and their geodesic distance is 1. Furthermore, it rigorously separates members inside a cohesive subgroup from outsiders. Because of this very strict requirement, large cliques are seldom found in real networks [16].

To address this rigid definition, the *n-clique* concept is introduced, in which the geodesic distance between any pair of nodes cannot exceed n and no node can be more than n links away from any others [15]. A larger value of n makes the clique more inclusive (with more nodes) but less cohesive (among its members). Another possible solution is the *k-core* concept, based on nodal degrees; in a k-core subgroup, each node is adjacent to at least k other nodes in the subgroup [16]. There have also been definitions of cohesive subgroups derived from the relative closeness of ties within the subgroup as well as the relative distance of ties from subgroup members to outsiders [16]. Readers who are interested are referred to reference [16] for more discussions and detailed explanations.

1.1.4 Structural equivalence

Many works on social network analysis focus on the network role and position analysis – that is, study of actors' structural similarities and patterns of relations. Two actors are perfectly *structurally equivalent* if they have exactly identical patterns of links sent to and received from all other actors [15]. That is, node i and j are equivalent if and only if the following conditions hold: (1) if node i receives a link from node k, then there is also a link from node k to node j; and (2) if node i sends a link to node k, node j also sends a link to node k. Structurally equivalence often causes fierce competition, as one actor can be easily replaced by another without affecting the network structure.

In real networks, the above definition is often too rigorous to be useful. A more practical scenario is that some nodes may be *approximately* structurally equivalent – that is, their relations to other nodes are similar but not identical. Many works have been done to measure the relation similarity between two nodes – for example, the Euclidean distance-based measurement, the correlation-based definition, automorphic and isomorphic equivalence, and regular equivalence [16].

Given these relation similarity measurements, the next step is to partition actors into subsets (also called *positions*), in which actors in one subset are closer to being equivalent than those in different subsets. There are many different ways to partition actors, including convergence of iterated correlations (CONCOR), hierarchical clustering, and multidimensional scaling. The last step is to describe the ties between and within positions – that is, how positions are related to each other. The commonly used methods – density tables, image matrices, reduced graphs, blockmodels, and relational algebras – have also been used for algebraic analysis of role systems. Details can be found in reference [16].

1.1.5 Other methods for network analysis

In the preceding sections, we focused on the study of "one-mode" networks linking actors to actors. *Affiliation networks*, also called *membership networks*, represent the involvement of a set of actors in a set of social events. An affiliation network is a "two-mode" network containing two types of nodes, actors, and events, and a set of relations between each nodal type. Research on affiliation network analysis aims to uncover the relational structures among actors through their joint involvement in events, and to reveal the relational structure of events attracting common participants [15].

A binary affiliation network can be represented using an *affiliation matrix* x, where $x_{ij} = 1$ if actor i participates in event j and $x_{ij} = 0$ otherwise. It can also be represented using a *bipartite graph*, where nodes are partitioned into two subsets, one including all the actors and the other with all the events, and one line in the graph links one actor to one event [16]. Galois lattices and correspondence analysis are often used to analyze affiliation networks; interested readers are referred to references [14,15] for detailed discussions.

In addition to the preceding analysis of deterministic (also called descriptive) social networks, probability and statistics have also been introduced in social network analysis [16]. Research topics in this area include statistical analysis of reciprocity and mutuality, structure inference, modeling and prediction for processes on network graphs, analysis of network flow data, and many others [16,19,20]. Readers who are interested are referred to references [16,19,20] for recent advances in this area.

Traditional social network analysis treats the network as a *static* graph, which is generated either from data aggregated over a long period of time or from data collected at a specific time instance. Such analysis ignores the temporal evolution of social networks and communities. To address this issue that has been overlooked, recently, there has been a growing trend to analyze how communities evolve over time in *dynamic* networks [21–24]. Lin *et al.* [25] proposed a unified framework to analyze communities and their evolution, in which the community structure provides evidence about how they evolve, and at the same time, the evolutionary history suggests which community structure is more appropriate.

1.2 Understanding media semantics in media-sharing networks

With the increasing popularity of media-sharing social networks, an important issue is to effectively manage these "billion-scale" social media and support accurate and efficient search of media objects. It requires accurate interpretation and understanding of media semantics. A promising approach is to annotate digital media with a set of keywords, such as "bridge" and "airplane" (sometimes called *labels*, *concepts*, or *tags*, in different contexts), to facilitate searching and browsing [26]. In this section, we first quickly review recent advances in social media annotation. We then focus on recent study on the emergent and evolutionary aspects of semantics and on leveraging social processes to understand media semantics.

1.2.1 Social media annotation

Depending on the flexibility of the keywords used to annotate digital media, media annotation methods can be classified into two categories, *labeling* and *tagging* [26]. In labeling, given a fixed concept set often called *ontology*, annotators decides whether a media object is relevant or irrelevant to a concept, whereas in tagging, which is ontology-free, users can freely choose a few keywords to annotate a media object.

The labeling process can be either manual or automatic. Manual labeling by users is tedious, labor-intensive, and time-consuming; Hua and Qi [27] proposed that the future trend for large-scale annotation is to leverage Internet users to contribute efforts. Ontology-driven automatic media annotation (also called *concept detection* or *high-level feature extraction*) and the closely related research area of content-based image/video retrieval have attracted much research activity in the past decades. These methods extract low-level content-based features that can be easily computed from digital media (for example, color histograms) and map them to high-level concepts that are meaningful and intuitive to humans. The mapping of low-level numerical features to high-level concepts (labels) is often done using learning algorithms – for example, neural networks, support vector machines, manifold learning, and feedback learning [26,28]. Recently, many automatic annotation systems have also used external information – for example, location information– to further improve the accuracy. Interested readers are referred to references [29,30] for review of recent works in this area.

Tagging enables users to freely choose the keywords (tags), and arguably provides better user experience [26]. Many social media websites, including Flickr and YouTube, have adopted this approach and encouraged users to provide tags to facilitate data management. In addition, the recent ESP Game motivates users to compete in annotating photos with freely chosen keywords in a gaming environment [31]. However, the free-form nature of tagging also poses new challenges: tags are often inaccurate, wrong, or ambiguous, and often may not reflect the content of the media [26]. To address the issue of "noisy" user-contributed tags, other context cues, such as time, geography-tags, and visual features, are fused with user-contributed tags to improve the search result and to recommend relevant tags [32]. Another approach to social tagging is to rank tags – for example, according to their relevance to the media content [33] and/or their clarity in content description [34] – to improve the visual search performance.

1.2.2 Semantic diversity

There are a few implicit assumptions in this concept detection framework – that is, the concept semantics is stable and the context is consistent. However, social media objects shared online originate from an unlimited number of sources [29], and there are an extraordinary large number of concepts that may not be shared universally [35]. The same keyword (concept) may have totally different meanings in different contexts. For example, on Flickr, the tag “yamagata” may refer to the Japanese town, the visual artist Hiro Yamagata, or the singer Rachel Yamagata [25]. Therefore, it is unrealistic to build one single classifier to learn all concepts for all media available on the networks, and it is of crucial importance to address this “semantic diversity” or “domain diversity” issue in media semantics in social networks [29,35].

To address this issue, some adaptive and cross-domain learning methods have been proposed, which efficiently and effectively adapt the concept detectors to new domains [36–38]. Zunjarwad *et al*. [39] proposed a framework that combines three forms of knowledge – global (feature-based distance), personal (tag co-occurrence probability), and social trust (finding people with correlated experience). The basic idea there is to use social trust and personal knowledge, and to recommend annotations only from people who share similar experience and opinions.

1.2.3 Emergent semantics

Social networks are highly dynamic and time-evolving, and so are media semantics. In real-world social networks, the visual representations of abstract concepts may change over time, and new concepts may emerge. In addition, some transient concepts that are relevant only to a specific event may exist for only a short period time [35]. In media-sharing social networks, it is important to consider and analyze the dynamic, emergent, and evolutionary aspects of media semantics owing to (explicit or implicit) collaborative activity, which is often ignored in media computing society [25]. Such investigation helps study how human beings interact with, consume, and share media data, and opens new views to the understanding of the relationship between digital media and human activities [25].

Emergent semantics has been studied in distributed cognition and sociology, and is defined by Cudré-Mauroux [40] as “a set of principles and techniques analyzing the evolution of decentralized semantic structures in large scale distributed information systems.” It not only addresses how semantics are represented, but also analyzes how self-organizing and distributed agents discover the proper representation of symbols (concepts) via active interaction among themselves [40]. Lin *et al*. [25] provided a review of recent works in emergent and evolutionary semantics in media-sharing social networks, interested readers are referred to reference [25] and the references therein for detailed discussion.

One challenging problem in emergent media semantics is *community discovery* – that is, how to extract human communities that collaborate on certain topics or activities [25]. For example, Flickr allows people to connect their images to communities via image "groups," in which images shared by a group of users are organized under a coherent theme [41]. But the challenge is to find the *right* community that will ensure *reachability* to other users for useful comments. Lin *et al*. [42] used the concept of "mutual awareness" to discover and model the dynamics of thematic communities. That is, users are aware of each other's presence through observable interaction (e.g., comments, hyperlinks, trackbacks), and the expansion of mutual awareness leads to community formation. The work of Lin and colleagues [43] extracted grammatical properties (the triplets of *people*, *actions*, and *media artifacts*) of interactions within communities, which will help generalize descriptors of communities. Lin *et al*. [44] analyzed the temporal evolving patterns of visual content and context in Flickr image groups, in an effort to understand the changing interest of users and to infer the genres of images shared in the group.

Another challenge is to characterize the *information flow* (or communication flow) in social networks and the roles that individuals play within the networks [25]. Such analyses are important in information source ranking and quality assessment, and identification of suitable time periods for marketing [25]. Choudhury and co-workers [45] proposed a temporal prediction framework to determine communication flow between members in a network, including the *intent to communicate* (the probability that one user wants to talk to another person) and the *communication delay* (the time taken to send a message). Their work showed that social context greatly affects information flow in social networks, in which *social context* refers to the patterns of participation (information roles) and the degree of overlap of friends between people (strength of ties). Choudhury *et al*. [46] developed a multiscale (individual, group, and community) characterization of communication dynamics; their analysis of the technology blogs (Engadget) showed that communication dynamics can be a strong predictor of future events in the stock market. Choudhury and colleagues [47] studied the temporal phenomenon *social synchrony*, in which a large number of users mimic a certain action over a period of time with sustained participation from early users, and a computational framework was proposed to predict synchrony of actions in online social networks.