INFLUENCE ANALYSIS: A SURVEY OF THE STATE-OF-THE-ART

MENG HAN*
Kennesaw State University
1100 South Marietta Pkwy, Marietta, GA, 30060, USA

YINGSHU LI
Georgia State University
25 Park place, Atlanta, GA, 30303, USA

(Communicated by Zhipeng Cai)

Abstract. Online social networks have seen an exponential growth in number of users and activities recently. The rapid proliferation of online social networks provides rich data and infinite possibilities for us to analyze and understand the complex inherent mechanism which governs the evolution of the new online world. This paper summarizes the state-of-art research results on social influence analysis in a broad sense. First, we review the development process of influence analysis in social networks based on several basic conceptions and features in a social aspect. Then the online social networks are discussed. After describing the classical models which simulate the influence spreading progress, we give a bird’s eye view of the up-to-date literatures on influence diffusion models and influence maximization approaches. Third, we present the applications including web services, marketing, and advertisement services which based on the influence analysis. At last, we point out the research challenges and opportunities in this area for both industry and academia reference.

1. Introduction. A set of social actors and a set of links among them construct a social network. The definition of a social network could be dated back to the late 1800s when both Emile Durkheim and Ferdinand Tonnies foresaw the phenomena of social groups [194]. The researchers in the fields of psychology, anthropology and mathematics work independently for the developments of social networks. From the definition by Rashotte et al. [195], influence, an important concept in social networks, is “the change from an individual’s thoughts, feeling, attitudes, and behaviors that results from interaction with other individual or group.” Influence is the natural product of information diffusion (or propagation) which is one of the fundamental processes taking place in social networks. Therefore, influence analysis occupies an important place in social networks.

Sociologists and other related scientists never stop trying to explore social networks since they also construct the modern social foundation. Many researchers have tried to test or examine that with whether there is an influence and how did people influence each other in social networks. And some results have been

2010 Mathematics Subject Classification. Primary: 68R10, 68W01; Secondary: 68W25.
Key words and phrases. Influence analysis, graph theory, network topology.
* Corresponding author: Meng Han.
achieved. Even so, prior to the Internet, quantitative data of social networks were scanty and the further influence analysis in social networks was in the slow-lane. In 2007, Nicholas et al. [47] published their years of research results based on the historical data from the spreading of obesity over 32 years. From the same research filed, David et al. [11] proposed another idea about the spread of obesity in social networks based on the simulations which further considered the group effect in obesity spreading. Both works tried to explore how the influence diffusion in social networks affects obesity. In their model, individuals’ influence over each other rely on food intake and physical activities [70]. Since other models consider obesity as a “contagious” phenomenon that can be caught if most social contracts are deemed obese, the interaction of social networks with environmental factors could not be explored. It was not accounted for in the general model where the social networks were proposed as a means to mitigate the obesity epidemic. Many other research results have been obtained recently, such as smoking behavior [48][52], happiness [68], and loneliness [25][117] which spread along a social network over time. At an American high school, Salath et al. [207] obtained high-resolution data of close proximity interactions during a typical day, and their work helps with the reconstruction of a social network for infectious disease transmission by using wireless sensor network technology [246, 98]. Through simulations, they showed that targeted immunization using the contact-network data is much more effective than random immunization. Stehle et al. [216] report a similar result like Salath [207] in a French primary school. The team headed by Stehle also provided several public-health implications of infectious diseases by collecting a period of history data from their experiments. By analyzing the real experiments in two middle schools in Germany, Ralf et al. [235] aim to test the operating social mechanisms that underlie the efficacy of bullying prevention programs. Kwon et al. analyzed how individual differences affect user’s intentions to use social network services by a Technology Acceptance Model (TAM) from psychology-based research.

Rosenquist et al. [201] tested the hypothesis that depressed people may influence each other from person to person in social networks. They also studied the effect from the structure of social networks in a psychiatric aspect. Customer churn prediction aims to detect customers with a high probability to attract. Based on two real life case studies via large scale data, Wouter et al. [232] found a significant impact of social networks on the performance of customer churn prediction model. By combining content-based and network-based approaches, Tang et al. [225] proposed some techniques to predict influence. Two medical data sets have further been tested to evaluate their proposed techniques called UserRank and Weighted in-degree. Based on Goyal and Kearns’ work in [78], He et al. [102] studied the Price of Anarchy of the competitive cascade game under the LT model in a theoretical aspect. Considering the price of a product in a social network, Francis et al. investigated the problem of how to find an optimal monopoly pricing and the relationship between the consumers and their neighbors. From a tie-strength perspective, Jichang et al. addressed the problems in social networks such as how fast does the information propagate, what is the role of weak ties for information diffusion, and so on. Zhao et al. [264] from another perspective gave some business suggestions for the cost-efficient and secured information propagation for online social networking sites such as pushing information to friends using a strong-tie-first strategy, and preventing privacy by removing positive weak ties from local communities. Rakesh Agrawal summarized the results of their recent investigations
around the nature of information, people and their relationships in social networks [2]. Their work includes information diffusion [3], analysis of opinion formation [19, 54], and factors influencing an individual’s continued relationship in a social group [24].

What the aforementioned results have in common is that they are from the real experiments based on the real social lives on questionnaires or laboratory tests, which are limited by the experimental size. These results are hard to be expanded to a large social entities. In addition, the methodologies mentioned above cannot be applied to large scale social networks [147].

Each month, more than 1.3 billion users are active on Facebook and 190 million unique visitors are active on Twitter. Furthermore, 48% of Facebook users who are 18-34 years old check their online page when they wake up, and 98% of 18-24 year old people are involved with at least one kind of social media. Online Social Networks (OSNs) have seen a rapid rise in the number of users and activities in the past years such as Facebook, Twitter, LinkedIn, etc., which means influence analysis in social networks has entered a new epoch. As an emerging part of social networks, OSNs represent most characteristics of traditional social networks in a digital version with a large scale. OSNs have kept growing for more than one decade and occupied an increasingly more important position in social networks. From OSNs, we can get more research results that were once unimaginable before. OSNs are not just a large continent size recreation or entertainment platform. Many OSNs could also be used for work purpose such as watching the market/competitors, and significantly and positively impact employees’ performance to some extent [136] [243].

The emergence of OSNs and the accompanying large amounts of data pose a number of both computational challenges and opportunities to academia and industry, especially those involving influence analysis. As far as we know, although OSNs have attracted a lot of attentions, limited works survey influence analysis. Bonchi [20], from a data mining perspective, summarized the applications around influence propagation in social networks. Guille et al. [82] gave a taxonomy result by dividing the main research challenges arising from information diffusion to three parts. Sun and Tang [217] from a computational aspect examined the research on social influence analysis. Their survey covers a lot of basic knowledge from a perspective of algorithm. However, in the past two years, new applications and algorithms have experienced an exponential development, thus a new comprehensive survey is extremely expected to give the overall reviews and guide the researchers in this area. The most recent survey on this topic is from Zhang [258], but they only consider the work of influence maximization itself without covering more literatures related to other parts of influence analysis. Fig.1 shows the publications regarding influence maximization in the recent years.²

In this survey, we focus on the latest problems and techniques regarding influence analysis in social networks. First, we give the bird’s eye view of the development of influence diffusion in traditional social networks and OSNs. Second, some preliminary knowledge regarding social networks including fundamental concepts of information diffusion and influence spread is presented. Third, we illustrate the most typical models which have already been widely applied for influence analysis. Then by analyzing the features and applicability of different models, we give the

¹http://www.statisticbrain.com/facebook-statistics/
²The statistic result focuses on data mining and social analysis and may not include all the literatures in the relevant areas.
comprehensive comparison. Forth, based on the literatures from different aspects, we point out some new challenges and opportunities in this new digital era, then propose a taxonomy which summarizes the state-of-arts. Finally, the newest applications based on influence analysis are introduced. We also put forth some future directions and possible improvements. In this survey, we not only provide a comprehensive analysis from the aspect of computer science, but also from other realms of academics such as precision science and sociology.

2. Preliminary knowledge related social influence. Graph, as one of the most important data structures, is an effective model to represent a social network. Given graph $G(V, E)$, where $V$ is the set of vertices (nodes) and $E$ is the set of edges (links), many features can be involved, such as properties of vertices, probability or weight of edges, etc. Fig.2 shows an example of the social network. In this graph, each node represents a person, and each edge represents the relationship between a pair of nodes such as friendship, colleague relationship or family relationship. We will give more details on how to measure influence together with the properties and features of social networks step by step [229].

2.1. Measurements in social networks. The two most important measurements in a social network are each node’s own properties representing user features and the relationships between them represented by edges. Influence of a social network is closely related to the nodes and the interactions among them [89].

2.1.1. Vertex strength. A node in a social network may represent a person, a group or an organization. The importance of a node is called vertex strength which could be measured by centrality indicating whether a node is a center node or key point of a network. The following metrics are adopted to measure centrality. Degree is the number of direct ties or connections that a node has [99]. In Fig.2, James has the highest degree centrality 6, the second highest one is Patricia whose degree centrality is 5. In terms of degree centrality, James is the most influential one. However, James can hardly influence the right clique, which can only be
reached through Patricia. Degree discount is a heuristic strategy proposed in [43] to improve the traditional definition of degree, where the degree of a node should be discounted if some of this node’s neighbors have already been selected as seeds.

Betweenness measures the number of times a node performs as a bridge in the shortest path between other two nodes [69]. Instead of James, Patricia is the most influential one if we consider betweenness centrality in Fig. 2. Patricia is the bridge for \{William, James, Mary\} and \{David, Richard\}.

Closeness evaluates how quickly a node can reach others in a network [99]. It computes the average shortest distance to all the other nodes. In Fig. 2, Patricia has the highest closeness centrality.

Similarity can be used to indicate two nodes have some common features, suggesting the probability that a node can influence another. The dotted line between Mary and Linda in Fig. 2 is an example meaning both of them are female and they graduated from the same university. Although they do not have a direct link, the similarity indicates the possibility for them to influence each other.

Eigenvector measures how influential a node is in a network. Compared with degree, it also considers the influence of a node’s neighbors rather than just the node itself. In Fig. 2, William has a higher eigenvector centrality than Elizabeth even though they have the same degree. PageRank [191] is a variant of Eigenvector centrality measurement.

2.1.2. Edge strength. Besides vertex strength, edge/tie strength measures on a pair of nodes which represents the basic influence processes and the interactions procedures between individual vertices. Generally, the edge strength can be classified into two groups, the direct relationship and indirect relationship. We then present the concepts and measurements around these two kinds of measurements.

1. Direct Relationship

In the social network, direct relationships could be considered as the link between two vertices. Most models use $G(V, E, p)$ to denote the graph of a social network, where $p$ represents the weight on the link which might be the influence between them or other relationship. One question came up, how does one evaluate the weight on a link? Due to the difficulties of finding the real information diffusion process, researchers have simply given several trivial solutions, such as assuming uniform probability for each link as the weight
(e.g., each link has probability \( p = 0.05 \), or the triviality model where the probabilities are selected uniformly at random from the set \( 0.1, 0.01, 0.001 \), or assuming the probability \( p(u, v) = 1/d_u \) where \( d_u \) is the degree of \( u \) \([124][42]\). Many learning models for the link weight evaluation have also been proposed. Saito et al. \([206]\) were the first to study how to learn the probabilities from a set of past diffusion history for the Independent Cascade model. They applied the Expectation Maximization (EM) to solve the problem they formalized. In the meantime, Goyal et al. \([71]\) proposed models and algorithms for learning the probability of influence in social networks. They used the Flickr data set which consisted of more than 40M edges and around 35M tuples actions to show their techniques have an excellent prediction performance. Different from Saito’s EM model which need to update the influence probability associated to each edge in every iteration, the solution of Goyal has a much better scalability. The term “interaction graphs” by Wilson \([242]\) has been proposed to impart meaning to online social links by quantifying user interactions. By their observations, several well-known social-based applications relying on graph properties have been verified. They also found that the use of real indicators of user interactions can get more realistic and more accurate results compare to basic social graphs.

2. Undirect Relationship

Besides the direct relationships in the network, the relationships between vertices in networks also potentially come from the undirect connections.

(a) Common Neighbor Based Relationship

The overlap of two nodes’ neighborhoods might decide the relationship between those two nodes \([80]\). Common neighbors are one of the most important features in the social network. It is not hard to have the feeling that two units in one network will have a stronger connection if they have more common neighbors. The common neighbor could be the same level unit, could have several common features. For common neighbor, the Jaccard distance is a very useful tool which can measures the similarity and the relationship between two nodes.

\[
JaccardDistance = \frac{|N(u) \cap N(v)|}{|N(u) \cup N(v)|}
\]

As the shown in Equation (1), where \( u, v \) denote the node in the network, and the \( N(u), N(v) \)denote the neighbor set of \( u \) and \( v \). The larger number of common neighbors they have, the closer their relationship will be. When two individuals are connected to many common friends, they are more likely to trust each other. On the other end, when nodes have no common neighbors, it is more difficult for them to trust each other.

(b) Reachability Based Relationship

Another undirect relationship can be described as reachability, which actually is an expanded version of common neighbor. The reachability of a node measures whether one node can reach another node, and how many different paths they can take to reach each other. Although the two nodes do not need have a direct common neighbor, after several hops they can reach the target vertex.

2.1.3. Uncertainty of measurements. In real social networks, the connections between users are not clear in different situations, especially when the data is incomplete, missing, or ambiguous. Uncertainty may be caused by components within
the network itself or by external factors that exist everywhere [247, 148, 118]. On one hand, most networks have constantly changing structures and features that remain dynamic [170]. For example, in a social network, a group of colleagues form a community when they are in the same company. In due time, such a colleague relationship may be broken as some of the members begin to work in another company while some of them start graduate studies. On the other hand, uncertainty is caused by the data generation process and the variety of networks. Different data acquisition techniques and data description methods may result in incomplete and inaccurate data which aggregates network uncertainty. Therefore, the process by which one identifies relationships in networks while considering uncertainty is very stringent.

It is difficult to account for the uncertainty among related nodes since traditional models do not make sense on uncertain networks [63, 62], and the inherent computational complexity of problems with uncertainty is always intractable. [96] investigates a framework for generating uncertain networks based on historical network snapshots. Four uncertainty construction models are presented to capture the uncertainty from dynamic snapshots, then the sampling techniques are employed to improve the efficiency of the algorithm. To describe the relationship of users in uncertain networks in a more practical way, the 2-hop expectation distance is adopted to approximate the expected number of common neighbors [97].

The number of common neighbors is one of the most important measurement for relationships among nodes. On one hand, common neighbors stand for direct relationships among nodes, since if an edge connects node $i$ and node $j$, they are also common neighbors of each other (each node’s neighbor set includes itself). On the other hand, the number of common neighbors also describes indirect relationships within a community. However, in an uncertain network, the concept of common neighbor is difficult to define since the direct relationship between a pair of nodes is not clear. Researchers use the expectation of an edge or path to measure a direct connection. Similarly, we use the expected number of common neighbors to represent the relationship.

In an uncertain graph $G$, the expected number of common neighbors between node James and Elizabeth can be calculated by the expectation of the number of distinct 2-hop paths between them.

In a deterministic graph, for node James and node Elizabeth, the number of common neighbors equals to the number of distinct 2-hop paths (distinct means any
two paths that do not have a common intermediate node) between them. As shown in Fig. 3, there are five nodes (Mary, Linda, Charles, Michael, and Barbara) between James and Elizabeth. Obviously, there are also four distinct 2-hop paths between them correspondingly. Apparently the number of distinct 2-hop paths and the number of common neighbors is a one-one correspondence. Then we can have a deterministic graph. For James and Elizabeth, a new distinct 2-hop path means adding a new node $v_k$ as a connector between them, and $v_k$ belongs to both James and Elizabeth, where Mary, Linda, Charles, Michael, and Barbara are the common neighbors of James and Elizabeth.

Obviously, in a deterministic graph, the number of common neighbors between two nodes corresponds to the number of 2-hop distinct paths between them. Since the expected number of common neighbors cannot be calculated directly, we use the number of 2-hop distinct paths to represent it. In an uncertain graph $G$, a 2-hop path is a one existing in some of the possible worlds generated from $G$. We cannot derive whether there is a 2-hop path or not; however, we can obtain the expected existence possibility of a path according to its existence situation in each possible world [96].

2.1.4. Other measurements. Since generally we only observe the times when particular nodes get infected but do not observe who infected them, Rodriguez et al. [197] tackled the problem that the underlying network over which the diffusions and propagations spread is actually unobserved in many applications. Thus, they developed a method for tracing paths of diffusion and influence through networks. The method employed time differences to infer edges, but there are many informative features such as textual content, which might give a more accurately estimation for the influence probabilities. Most recently, Zhang and Tang, et al. [260] proposed a new metrics to measure the relationship of two nodes in network based on random path similarity. An search algorithm named “Panther” was proposed to efficiently answer the top-$k$ similarity query. Considering that influence is diffusing according to path connection in social network, “Panther” is a very efficient and scalable approach to be applied in influence analysis.

In addition to this, when two nodes connected by one edge do not share a common neighbor, it is hard to well explain the observed edge sign. [214] addressed this problem by applying a new model for different node types. Initially, the authors analyzed the local node structure in a fully observed signed directed network, inferring underlying node types. They proposed that the sign of an edge between two nodes must be consistent with their types, and this result could explain the edge signs even without the common neighbors between two nodes.

2.2. Structure and properties of social networks. All the above measurements give us a way to realize social network from a microcosmic angle, further the structure and global properties of network are going to be introduced in this subsection to provide a relative macrocosmic point of view to comprehend social networks. The structure features and properties of social networks could lead to further profiling of influence diffusion in social networks. From both microscopic and macroscopic aspects to study social networks could allow us better understand users’ behavior and actions then develop a more thorough understanding of relevant influence research results.

2.2.1. Structure of social networks. In social networks, two users construct social tie, three users compose social triad, and users more than three could build clusters
or communities. These kinds of structure features above are very common, and all
these phenomena are due to the social nature of human beings. From aspect of
influence diffusion, if more than two users belong to same group or community, they
are very likely to be friend, and easily to be influenced by actions or behaviors from
each other. Several famous reports such as power law theory and small world theory,
et al. have been proposed to describe the structure features of social networks. Due
to space reason, we only give brief introduction to related topics as following:

1. Power Law. As introduced in previous section, degree of a node is the num-
ber of edges connected to that nodes. Power law in statistics is a functional
relationship between two variables, where one variable will be changed ac-
cording to the variation of another. Let function $P(k)$ be the probability of
one random node with degree $k$. The plot of $P(k)$ for the whole network that
generates a histogram of degree distribution of nodes is similar to a long right
tail. The long right tail indicates that a small proportion of nodes have a very
high degree while most nodes have a low degree. Social networks keep follow
the long right tail of power law theory have been studied and been verified fairly true [130][178][140].

Some online social network also reported statistic result to support power
law. Twitter, for example, the top twitter users such as Katy Perry has
75,375,552 followers, and Justin Bieber has 65,523,692 followers, but the
average follower per user is just 208. Similar situation is also applicable to
Instagram and LinkedIn.

2. Small World. Based on empirical study of social networks, Michael Gurevich
conducted his result in 1961 and later he concluded that “it is practically
certain that any two individuals can contact one another by means of at
most two intermediaries” [56]. Milgram continued Gurevich’s experiments in
acquaintanceship networks and published their famous paper “The small world
problem” [176]. Both research results were done before the era of internet and
thus limited by the sample space.

Watts et al. published their paper [240] considering the “small world”
problem and one corresponding model to generate random “small world” net-
works. Until 2003, Columbia University conducted one analogous experiment
on social network amongst Internet email users. Their effort included 24,163
e-mail chains, for 18 targets from 13 different countries all around the world
and involved 100,000 individuals [59]. Among the successful chains, it is more
common that shorter lengths only reached their target after 7,8 or 9 steps.

Recently, Facebook3, for example, reported that the average number of
Facebook friends for US females is 250, which is many times larger social
network if compare to e-mail network in 2003. Based on all result above,
social network is much more smaller that most of us imaged especially in our
internet era. At the same time, influences in social network among each other
are spreading in day and night in the whole world.

How information propagates through the social network has been studied for a
long time. Several potential models have been proposed to capture the structure
of the social networks. Considering topics in social networks, Gruhl et al. [81]
pointed out that the popularity of different topics might remain constant in time
or become more volatile. Kumar et al., based on their previous work [130] which

---

3https://www.facebook.com/
analyzed community-level behavior of users, proposed that much of the behaviors were characterized by the stars type propagating model [131]. A game-theoretic framework was introduced to address the community detection problem based on the structures of social networks in [39] by Chen et al. Since the organization of the network plays an important role in the social networks, Arun et al. [172] proposed a method to infer social hierarchy. Focusing on blog networks, Leskovec et al. [140] proposed a epidemiological model to capture the topological characteristics of social networks. Based on the analysis result, [140] also reported that most topological network characteristics follow power laws which include in-degree, out-degree, and cascade size etc.

Recently, a series of results focusing on triadic structure in social networks have been proposed by Tang’s group [163][113]. Different from community, triadic represents the three nodes unit, which is very common as well as pairs in the social network but with more interesting features.

The structure of social network decide the mode of influence propagation to some extent, and give researchers more room and opportunities to develop valued applications such as advertisement services, recommendation systems, etc.

2.2.2. Properties of social networks. Besides structure features of social networks, some basic network properties such as: size of network, represented by the number of edges in the network; order of network, considers the number of nodes in the network; and density of network, which is applied as a measure of network health and effectiveness, etc. also reflect their own effects in influence analysis.

In the early years, Xu, Yuruk, et al. [245] proposed “Scan”, a structural clustering algorithm, to find out clusters, hubs and outliers in network. Considering the different structure properties, “Scan” assigns two vertices in the network to a cluster, identify hubs, and recognize outliers according to how they share neighbors. Unfortunately, one of weakness of “Scan” requires high computation costs for large-scale graphs because before identifying hubs and outliers, “Scan” has to find all densely connected node sets as clusters. Just recently, Shiokawa, Fujiwara and Onizuka [210] from Japan proposed a improved version of “Scan” named “Scan++”, which could detect the same clusters, hubs, and outliers as same as “Scan” but with much shorter computation time.

Tiancheng et al. [163] studied how links are formed in social networks and especially focused on investigating how a two-way link formed. The leaning framework they proposed formulated the problems of predicting reciprocity and triadic closure. Structure hole have been verified by a few empirical studies. Tiancheng et al. [162] defined the problem of mining top-k structural hole spanners and provided a quality function to formalize the problem. Their studies show more evidence for the theory of structural holes such as how detecting structural holes spanners can help other social network applications to do the kernel detection and link predication. Bakshy et al. [12] from Facebook conducted two very large field experiments that identify the effect of social cues on consumer responses to advertisement, measured in terms of ad clicks and the formation of connections with the advertised entity. And the result from their experiments has a guide significance for advertising optimization, user interface design and other analysis in social science research.

2.3. Remarks. In this section, we give the necessary preliminary knowledge regarding to social network, kinds of measurement, and properties effect of the result in terms of influence diffusion among the network to a great extent.
Although more attentions have been paid to the analysis of the interior structure, relationships and macroscopic analysis in social networks, there are still many exciting directions to pursue around the understanding of social networks such as structure dynamics of networks, community detections, and the properties of heterogeneous networks.

3. **Influence analysis.** We can divide the influence models based on the statistical scope as shown in Figure 4. There are two categories of models if considering the social network characteristics. The first kind of model can be named as static influence models, which are simple and easy to assess. These kind of models have been developed in different aspects, where it is assumed that the influence between each node is static and time independent. Another group of models are named dynamic influence models, which allow the influence change over time. We will address more research results around the dynamics and evolution in the following sections. Generally, the use of the snapshot to capture the dynamics of a network is a very intuitive method, another universal solution for dynamic networks is the building of evolution process or distribution by time stamp which can obtain the changes from the network [209].

Static influence models generally are used to find or select the most influential nodes at that moment. The static network is fixed in both the size and topological aspects. Most of the influence models based on static networks that have been proposed have also kept the influence between nodes stationary. Several metrics such as degree distribution of nodes and structural features of network have been proposed and utilized as the measurement to maximize the influence.

In the work of Habiba et al. [134], they first extended standard structural network measures to dynamic networks, then ranked the blocking ability of individuals by the new dynamic measure. Based on their analysis, key blockers in a network can be identified by their simple, practical and locally computable algorithms [255]. In 2010, scholars from Harvard University harnessed data on Facebook applications to study the role of social influence on the dynamics of popularity. By tracking the popularity of a complete set of applications installed by the users in Facebook, they captured the behavior of all individuals who could influence each other in that context [190]. Viswanath, Mislove, et al. [234] studied the evolution of activity between users in the Facebook social network to capture the fact that over time social links can grow stronger or weaker. [234] also reported that links in the activity network tend to come and go rapidly over time, and the strength of ties exhibits a general decreasing trend of activity as the social network link ages. Other researchers found that people who shared information about similar types of music...
and movies (but not books) were more likely to be friend one another by analyzing the Facebook activity data from a group of college students over 4 years from another perspective [142]. Cha et al. [31] characterized how information spreads over current online social networks. They collect and analyzed data from the Flickr, which involves 2.5 million users and 11 million photos. Rossi et al. [202] proposed a temporal behavior model that captures the “roles” of nodes in the social network and how they evolve over time.

Influence analysis in dynamic networks has been a very active research area recently. Based on a small set of “snapshot” observations of a social network and detailed temporal dynamics, Dan et al. studied the relationship between these two ways of measuring influence [51]. In [122], Kempe et al. presented a model of cultural dynamics that captures the aspect of the interplay between selection and influence. Chen et al. [192] proposed an influence model which incorporated dynamic parameters to learn how influence changes over time. Three examples were provided to show the practicality of their model. Fan and Shelton [66] provided a sampling-based learning algorithm for modeling the continuous-time social network. Zhuang et al. [268] considered the changing over the network, and aimed at probing a subset of nodes in a social network to approximate the actual influence diffusion process.

Another work regarding to exploring and predicting information diffusion in temporal dynamic network was developed by Bourigault, Lagnier, et al. [23]. From a learning aspect, the information diffusion processes is learned by embedding users in a continuous latent space, and this strategy bases on the information content that allow the algorithm learn a threshold to split users in one contaminated group and one non-contaminated group.

Although the above works take into account the dynamics of social networks, there seems to be very limited understanding of the inherent dynamic properties of social networks, and the most of them did not involved the real applications based on the dynamic of network’s structure.

Authors of [196] modeled diffusion processes as discrete networks of continuous temporal processes occurring at different rates. They created a model that presents a method for inferring the mechanisms underlying diffusion processes based on observed infections. However, since their model is based on some assumptions to the spatiotemporal structures that generate diffusion processes, it is hard to employ the model directly to the real world for many applications.

To analyze the dynamic of Twitter, Mayers and Leskovec [182] studied ways in which network structure reacts to users posting and sharing content. They found that information diffusion in the form of cascades of post re-sharing often creates sudden bursts of new connections, which significantly change users’ local network structure. They also propose a model that quantifies the dynamics of the network and the occurrence of these bursts as a function of the information on spreading through the network.

The major diffusion models of social influence is shown in Figure 5. Leskovec et al. [139] modeled the outbreak detection problem, and proved that the influence maximization problem was a special case of their new problem. A “Cost-Effective Lazy Forward” (CELF) scheme has been proposed which uses the submodular property to achieve 700 times speedup in selecting seed vertices compared to the basic greedy algorithm from Kempe et al. [124]. As discussed in Chen et al. [42], CELF

https://www.flickr.com/photos/tags/flickr/
still faces serious scalability problem. Chen et al. proposed new heuristics based on the arborescence which could handle graph with million of nodes and have influence spreads close to the greedy algorithm while at more than six orders of magnitude faster than the greedy one. Recently, Leskovec et al. also proposed the extended version of CELF named CELF++ [76] which tackles the shortcomings of CELF, and they reported that CELF++ is 35-55% faster than CELF.

Another greedy algorithm named SMG which stands for State-Machine Greedy was proposed recently by M. Heidari et al. [107]. The main idea improves the speed of greedy algorithms by preventing recalculation done by older methods. SMG improved upon the traditional greedy algorithm from a time complexity standpoint by triggering nodes in the startup queue, reducing time of graph construction and preventing re-traversing of nodes. From their experiment, SMG has a much better performance than CELF. However, their paper does not concern the effect of structure on the time complexity which still has an open problem in this kind of research.

In economics, Luca, et al. presented an experimental investigation of persuasion bias. And they found that the social influence not only depends on being listened to by others, but also on listening to many others. They investigated how the communication structure of a social network affects the aggregation process then how to determine the social influence [50]. Data from a nationally representative US sample was analyzed to determine whether and how social ties related to behaviors that determine a household’s carbon footprint. By adopting a probability-based approach to measure distinct profiles of social relationships, two dimensions of social relationships, norms and strength of ties are considered in their work [233].

In decision science, similar to game theory, which is between concerned with identifying the values, uncertainties and other issues relevant in a given decision, its rationality, and the resulting optimal decision, the famous Hoede-Bakker index computes the overall decisional ability of a player in a social network, but the main drawback is it hides the actual role of influence function, analyzing only the final decision in terms of success and failure [108]. Michel et al. separate the influence part from the group decision part and focus on the description and analysis of the influence. In the original Hoede-Bakker index, a set of all players which includes agents, actors, voters denoted by \( N := 1, \ldots, n \) is contained in a social network. The players need to make a certain acceptance-rejection decision. Each player will either to say YES or NO by an inclination vector denoted by \( i \) which is a \( n \)-vector consisting of ones and minus ones. Assume players may influence each others and due to the influences in the network, the final decision of a player may be different from his original inclination. The final influence result is decided by the vector \( i \) and a group decision function \( gd(Bi) \) [79]. Toni et al. proposed a paper published by Oxford University Press on behalf of the Gerontological Society of America given a
convoy model to explain the social relations from a multidisciplinary perspective [8]. Lso, Fond and Neville [133] measured the gain in correlation and assessed whether a significant portion of this gain is due to influence and/or homophily for temporal network data where the attributes and links change over time.

From the aspect of opinion dynamics, Das, Gollapudi, et al. [53] considered the problem of modeling how users update opinions based on their neighbors’ opinions. Essentially, the opinions changing based on neighbors’ opinion is the influence from neighbors. A set of online user studies based on the celebrated conformity experiments of Asch [9] are performed. The authors of [53] showed that most existing and widely studied theoretical models do not explain the entire gamut of experimental observations, and consensus and polarization of opinions arise naturally in their model under easy to interpret initial conditions on the network.

It is even more difficult to pursue the optimal node sets that can maximize the influence in a dynamic social network. Besides all the other challenges, updating a network to reflect its dynamic nature with time is extremely resource consuming in large social networks. Therefore, [95] proposed an efficient integrated solution to select the most influential nodes in dynamic social networks considering the challenges and features of OSNs. In addition, the model BICOT could control the balance between influence depth and breadth. It is the first step to explore the potential of broad influence maximization. Through comprehensive experiments tests, the results show that ICOT model can achieve a comparable influence diffusion result to the learning-based algorithm but does not need the strict input requirement; and at the same time, has a much broader influence coverage.

Fig.6 shows an example of the influence diffusion in a dynamic social network. The top figure presents the network at time $t = 0$; the middle figure describes the changing of the network at time $t = 1$; and the bottom figure is the network topology at the end $t = 2$. From left to right, the network is divided into three communities. Through the time flow from top to bottom, we could get that only the two communities (the left and the right in the dashed line frames) have changed their topology. But the nodes in the middle part remain unchanged. From this example, we notice that it would be more efficient if we could identify the two changed communities but ignore the middle community during updating. Therefore, probing the most active communities to approximate the global evolution of the network would be a very effective solution. Most existing research surprisingly ignores the advantages of the community feature. On the other hand, the complexity of solving the influence maximization problem rapidly increases with the size of the network. Therefore, finding influential local nodes in each relatively smaller community could be much more efficient.

It’s worth pointing out that the objective of our work is to track the network’s global dynamics as well as to reduce the cost brought by frequently updating the whole network. We utilize the “community” instead of “node” as a unit to probe the change of the network because the community is the basic and natural structure in large networks, which is a better choice compared with the node. Even though the updating unit is the community, the changing of nodes and links among nodes in the communities are more commonly the updating targets. For each iteration, based on our theoretical analysis, when $b$ communities are selected to be actually updated, the nodes and the links among nodes in the selected communities are going to be updated. The reason we do not take a node as the unit to update the network is: from an overall perspective, even the changing of only one node
in a network will only result in the changing of several relationships. Frequently updating the network node by node will bring in more redundant costs because of updating their neighbors. On the other hand, a community will cover several nodes with closer relationships. The observations in previous results [268] show that most of the dynamics in large networks have some kind of local effect [202] confirming the advantage of communities over nodes. The local update of communities could update the dynamic changing within specific areas.

Again, considering the node “Michael” in the left frame, from time $t = 0$ to $t = 2$, “Michael” disappears from the original topology. In this case, we could consider it as node “Michael” has been excluded from the network at time $t = 0$, which is
the opposite process of a new node joining. Our algorithm considers the changing
of each node within the selected communities. Both new nodes joining and leaving
are considered in the algorithm. Thus, our algorithm considers the community as
a unit to probe the dynamic of networks from a global sense but the dynamics of
nodes is the actual cell being updated.

Although it is not always the case that one influential node in a community
corresponds to an influential node in the whole network, apparently an influential
node in one community has a stronger influence based on its degree and neighbors’
density compared with normal nodes in the whole network.

3.1. Learn probability from social network. One of the basic questions of in-
fluence analysis is how to gather the data of social networks, and how to evaluate
the relationship or transmission routes between the entities from each other. Most
of influence maximization problems assume that the social network structure and
influence probabilities have been given as input. A precise structure of the network,
and applicable influence probabilities have a substantial impact on the problem’s
final result. However, it is a non-trivial work to extract a social network’s structure
and compute the probabilities precisely between each other. There are many pos-
sible relationships implicated in social networks and different relationships might
correspond to different influential probability. Several efforts have been made to-
ward correcting these issues.

To analyze influence, the first problem is understanding the relations in social
data. But in real life, uncertainty exists in all kinds of networks. The uncertainty
may result from network components themselves or external factors. How to figure
the uncertainty of influence in social data out is the very first challenge. How-
ever, in practice, a clear relationship among pairs of nodes is difficult to detect in
huge uncertain complicated networks. Due to the increase of complexity in mod-
ern networks especially social networks (Facebook, Twitter, and LinkedIn, etc.), it
becomes more and more difficult to efficiently identify the relationship in networks.
Considering the uncertainty of social network, [97] designed a method for relation-
ship detection in uncertain networks. The entities in a same community or group
with relationship usually interact frequently, share similar properties and generate
common features. A two-hop expectation distance was adopted to approximate the
expected number of common neighbors. This method can also serve as a framework
for measuring the expected number of common neighbors in uncertain graphs.

Anagnostopoulos et al. [6] categorized three types of reasons for the correlations
in social networks. The first one is influence where the action of a user triggered
his/her friend’s recent actions, the second one is homophyly which means similar
individuals often perform similar actions, and the third one is environment where
external factors are correlated both with the relationship of two friends and their
actions. Gruhl et al. [81] first derived an Expectation-maximization (EM)-like algo-
rithm by using a variant of the independent cascade models to induce the influence
probabilities. More formally, Saito et al. [206] studied how to learn the probabilities
from a set of past diffusion history for the IC model. They apply the similar
EM to solve the problems they formalized. Individually, at the same time, Goyal
et al. [71] built models of influence from a social graph and the log of actions by
the users belonging to the network. By introducing the credit distribution, Goyal
et al. [72] proposed a framework to maximize the influence as well as learning real
influence from users’ history log data at the same time. Consider a situation in
which information can reach a node via the links of the social network or through
Influence Analysis

Influence Analysis in Dynamic Network
Learn Probability from Social Network

Figure 7. Influence Maximization Models in Social Network

the influence of external sources such as newspapers, TV stations, and online news sites. Myers et al. [183] proposed a model in which information can reach a node via the links of the social network or through the influence of external sources. By using a one month trace of Twitter, they studied how information reaches the nodes of the network. They quantified the external influences over time and found how these influences affect the information adoption.

Besides the modeling of influence maximization, some general frameworks with influence learning are also proposed recently. [137] consider the influence maximization on influence probability in the absence of complete information. Online influence maximization (OIM) [137] is proposed which tried to figure out the problem that learning influence probabilities as well as running influence campaigns at the same time. Different from [72], OIM has some existing influence information at the beginning, then by adopting the Explore-Exploit strategy, the model can select seed nodes using either the current influence probability estimation (exploit), or the confidence bound on the estimation (explore). The framework OIM can be used to most of exiting IM models since OIM actually provide a mechanism to optimize the process of influence maximization.

How to evaluate the real interaction and influence in social networks is one of the most important original problems for influence analysis. However, this still remains largely unexplored due to the complexity of relationships and structures within social networks.

4. Influence maximization in social networks. Each month, more than 1.3 billions users are active in Facebook and 190 million unique visitors are active on Twitter site. Furthermore, 48% of 18-34 year old Facebook users check their online page when they wake up, and 98% of 18-24 year old people involve at least one kind of social media 5. With the high percentage usage, OSNs have become one of the best effective and efficient solutions for marketing and advertising. The basic problem of influence maximization can be described as follows: in a social network which include nodes and edges, all the nodes have influence between each other. Try to select an initial set of \( k \) nodes such that they eventually influence the maximum

\(^5\text{http://www.statisticbrain.com/facebook-statistics/}\)
other nodes based on some kind of models. Figure 7 is the basic category of different influence maximization models in social network.

Domingos et al. [60] introduced the problem of identifying influential customers in marketing campaigns as a learning problem. Then, in 2003 Kempe et al. [124] studied the influence maximization problem for two fundamental information diffusion models, which is the independent cascade (IC) model and the linear threshold (LT) model.

In both of the two models Kempe introduced, the input is a network with nodes and edges, where each node is either active or inactive, and the possibility of one node becoming active increases monotonically as its neighbors become active. If one node becomes active, it will never be inactive again. How to maximize the influence in social networks depends on their influence model. Therefore, we present the classical models by different categories.

4.1. Cascading model. In the IC model, at the beginning of time \( t_0 \), nodes which are active are much similar to some “seeds” in the network and these nodes are considered contagious. One node \( u \) has one chance of influencing each inactive neighbor \( v \) with probability \( p_{u,v} \) which can be considered as the ability of the influence from \( u \) to \( v \). If this attempt succeeds, node \( v \) becomes active at time \( t_1 \). This process iterates and continues until no new node becomes active in the network. [228] studied the strategies selecting seed users in an adaptive manner to maximize the influence in social network. A Dynamic Independent Cascade (DIC) model based on IC is proposed to capture the dynamic aspects of real social networks. Hu, Meng, et al. [110] focused on the IC model and proposed a series-parallel graph based approach to improve the efficient and accurate with a linear time complexity.

The Decreasing Cascading (DC) model [123] tried to reflect the information saturation problem more practically. In DC model, the probability of activating a node will decrease if the attempts have been made by more people.

4.2. Threshold model. In the LT model, it is the same initialization as IC, each node \( v \) will be influenced by all \( \{u_1, u_2, \ldots, u_i\} \) from \( v \)'s neighbor set \( N(v) \) according to the sum of the weights of \( \sum_{i=1}^{N(v)} p_{u_i,v} \), such that the sum of all the incoming weights to \( v \) is less or equals to 1. The node \( v \) chooses a random threshold \( \theta_v \) uniformly from \([0, 1]\) at each time stamp. If the sum of weights from all the active neighbors of an inactive node \( v \) is more than \( \theta_v \), then \( v \) becomes active at the next time stamp. This process also repeats to the end until no new node becomes active. Kempe et al. first formulated this problem as a discrete optimization problem in [124]. Considering a social network as a graph \( G = (V, E) \), where \( V \) and \( E \) is the set of vertices and edges with size \(|V|\) and \(|E|\). Choose an influence diffusion model (IC or LT) and an initial active seed set \( S \subseteq V \), the expectation of the active node’s number at the end of the process is the expected diffusion spread of \( S \), denoted as \( \delta_m S \). Then the influence maximization problem is defined as follows: To find the best seed set \( S \) to maximize the \( \delta_m S \) in a directed social graph \( G = (V, E, p) \) where \( p : E \to (0, 1) \) is the function assigning each edge \( e \in E \) a probability \( p(e) \).

Chen et al. [42] has proved the problem of computing the expected influence spread \( EIS \) of node is \( \#P \)-hard. Under the LT model, Lu, Fan, et al. [168] investigate influence spread estimation for influence maximization in an efficient way. In [168], the authors show that the \( EIS \) of a node could be computed by finding cycles through it, and they also developed a more efficient approximation algorithm to solve the problem.
4.3. Voter model. Targeting to select the best seed nodes set, the Voter model introduced by [49] has also been invested by several literatures. In the Voter mode, each node is influenced with probability that is proportional to the number of neighbors which were influenced already [180]. The voter model states the property in social network that a person is more likely to keep or change his/her opinion to the direction held by most of his/her neighbors. Different from threshold models which is monotone in the sense where once a user becomes “activated”, then (s)he stays activated forever. The voter model is suitable for some cases such as which opinion a user is currently hold but could be changed later, meaning not monotone all the time [65]. Formally, assuming a set of nodes has been activated at time $t_0$, in the next time slot $t_1$, the probability of one inactive node $u$ could be influenced is $p_u = \frac{|N_a(u)|}{|N(u)|}$, where $N_a(u)$denotes the activated nodes in $u$’s neighbor set, and $N(u)$ is the whole neighbor set of $u$.

An voting patterns application in online content is proposed by Sipos, Ghosh, et al. [213]. They explore how users respond to question such as “Was this content helpful?”. By using the data from Amazon product reviews, they show the relationship among the independent voting decisions actually are influenced from each other and based on the context.

Different from the models we introduced above, another kinds of models which based on voting are introduced by scholars Wang et al. [237] proposed Positive Influence Dominating Set (PIDS) selection algorithm to find the seeds set to influence the network. The basic idea is if more than half neighbors of an individual have positive impact on him/her, then the probability that this individual’s positively impact on others will be high. And the influence diffusion iterates on this process until no new active nodes appeared.

4.4. Time constrained model. Goyal, Bonchi et al. [71] have already shown that time plays an important role in the influence spread from one user to another and the influence is also different because of various relationships between users. Liu, Cong, et al. [156], proposed a time constrained influence maximization model. After showing the NP-hard complexity of the problem, they generalize their proposed algorithms for the conventional influence maximization problem without time constraints which could be utilized in other similar problem.

4.5. Budget allocation model. In [74], Goyal studied the alternative optimization problems which motivated by another two constraints from the classical model. In the basic IC and LT models, the input include a network $G$ and a seed size parameter $k$. And the objective is to maximize the expectation of influence. Alternatively, in the study of [74], the authors tried to optimize different targets such as the size of seed set and the influence time. Different from the IC and LT, in their first model a threshold of the influence expectation is given as input, in this sense a smaller the proactive seed set means a smaller budget of the process, and they provided greedy algorithm to solve the models they build.

With respect to allocation, Hatano, Fukunaga, et al. [100] considered the influence maximization with three participants: advertisers, customers, and publishers into play. The purpose of advertisers is to maximize the influence on customer decision and convert potential customers into loyal buyers, subject to budget constraints. To overcome the substantial computational cost, the authors proposed a algorithm based on Lagrangian decomposition. The key idea of Lagrangian decomposition is to decompose the optimizing problem into several subproblems by
introducing auxiliary variables and a Lagrangian relaxation of the problem. Because
the objective function and the capacity constraints share no common variables in
the relaxation version problem, it is possible to decompose the problem into sub-
problems, for which greedy algorithms will perform well.

4.6. Competitive diffusion model.

4.6.1. Bilateral competition diffusion model. Bilateral competition diffusion model
could be considered as there are two opposite opinions in the social scenario, where
one could be positive and another is negative. How to analyze the diffusion process
is a very challenging and meaningful problem. In real life, it is often the case that
different and often opposite information or ideas are competing for their influence
in the social networks. Such competing diffusion could range from two competing
companies, two political candidates of the opposing parties to even the government
tries to inject truth information to fight with rumors spread in the public.

Goyal et al. [73] gave more approximation analysis of influence spread based on
models they developed in [74] which considered the alternative algorithm goal of
influence maximization. Li et al. [151] extended the classic voter model to signed
networks and analyzed the dynamics of influence diffusion of positive and negative
which represent two opposite opinions. For short term and long term dynamics,
they derived the exact and closed-form formulas separately. He et al. [104] studied
the problem that one entity tries to block the influence propagation of its competing
entity as much as possible by strategically selecting a number of seed nodes. They
model the competitive linear threshold (CLT) as an extension to the classic LT
model. Consider the situation that one company wants to popularize a new product
where a competing product is already being introduced. Carnes et al. [29] propose
two models for the simultaneous diffusion of two competing technologies on any
network which reduce to the independent cascade model of Kempe et al. The
“follower” in [29] is the player who selects seed nodes with the knowledge that some
nodes have already been selected by its opponent.

Recognizing that companies are competing in a viral marketing, Lin and Liu
[154] formulate the competitive influence maximization in a “General Competitive
Independent Cascade (GCIC)” model. GCIC also describes the general influence
propagation of two competing sources in the same network.

Considering the PIDS selection problem, Wang et al. [237] propose the influence
maximization algorithm based on the idea that as more neighbors of an individual
have positive impact on one user X, the positive impact from X will be higher.
Another work focusing on positive influence in online social networks is proposed
by Zhang, Dinh, et al. [257]. They proposed a two-phase model called Opinion-
based Cascading (OC), which also has a NP-hard complexity and impossible to
design any approximation algorithm with finite ratio unless \( P = NP \).

Tsai et al. [231] consider the situation where two parties have to make their
choices without the opponents’ choices in competitive diffusion networks. Similar to
[231], considering the competition among similar products or services from different
companies, Lin et al. [153] proposed a data-driven model STORM to maximize
the expected influence in the long run. Most of earlier works are based on model
driven methods [231] [165] [18], which apply specific heuristic to choose the seed
nodes in the network given a known influence propagation model (e.g. IC or LT).
STORM is capable of learning a good multi-party influence maximization strategy
which utilizes arbitrary existing single-player influence maximization strategies as its actions, and finds the best policy to select them given the observed conditions.

4.6.2. Multilateral competition diffusion model. Borodin et al. [22] gave several natural extensions to the linear-threshold model named $K$-$LT$ and provided the algorithm which the original greedy algorithm cannot work. $K$-$LT$ reflects several phenomena of competitive influence propagation that match our daily experience. Similarly, from the perspective of the the owner of the social network platform, influence in competitive viral marketing was considered by Chen et al. [165]. They proved that the fair seed allocation is NP-hard, and with the properties of monotonicity they developed greedy approximate algorithm to solve their problem. Different from $K$-$LT$, the model in [165] considers the phenomena that influence decays very quickly in time, and customers are more likely to rely on recent information than on old one.

As shown in Fig. 8, two nodes have been selected as the seeds which marked as active in the social network, then the active nodes try to influence their neighbors by a probability. If the neighbor was influenced, then the status turn from inactive to active, and continue to repeat the process, as shown in Fig. 9. When one active failed to active one neighbor, it will not try to influence the neighbor any more. The whole process stop until no new active node generate.

The $IC$ and $LT$ model together with their extensions set the foundation of most existing algorithms to maximize the influence in OSNs.

For both $IC$ and $LT$, Masahiro et al. [127] has achieved a good reduction in computational cost by estimating all the marginal influence degrees of a given set of nodes on the basis of bond percolation and graph theory. Previous greedy algorithms still face serious scalability problem. Chen et al. [42] showed that computing influence spread in the independent cascade model is $\#P$-hard problem. To address the scalability issue, they proposed efficiency heuristic algorithm by restricting computations on the local influence regions of nodes. Additional, a tunable parameter for users to control the balance between the running time and the influence spread of the influence. Addressed the scalability also, heuristic algorithm designed by Wang et al. [236] can be easily scalable to millions of nodes and edges in their experiments under $IC$ model. A power-law exponent supervised Monte Carlo method is utilized to efficiently estimate the influence spread for nodes with specified precision by sampling only part of child nodes [161].

A mixed integer programming (MIP) formulation with elements from stochastic optimization and network design were introduced by William [208] to maximize
the expected spread of cascades in networks. Different from the classical model, William’s model is more general to capture adding edges, or to increase the local probability of propagating the cascade. In such situation, the objective function to maximize the influence no longer submodular. They contribute a set of preprocessing techniques to reduce computation time for their algorithms.

In [121], the authors proposed algorithm IRIE where IR for influence ranking, and IE for influence maximization in both classical IC model and the extension IC-N model incorporating negative opinions [37]. [121] reported that their algorithms scale better than PMIA [42] with up to two orders of magnitude speedup and significant savings in memory usage, while maintaining the same or even better influence spread. For LT model, Chen et al. [44] show the #P-hardness by using the interpolation technique, which is much harder than the reduction in [42]. They also showed that the influence computing in directed acyclic graphs (DAGs) can be done in linear time, and based on this result a scalable heuristic algorithm were developed tailored for the influence maximization in the LT model. Since Chen’s model [44] relies heavily on finding a high quality LDAG which is also NP-hard, the heuristic algorithm have to be used introduce an additional level of loss in quality and the memory cost is expensive. Goyal et al. [77] addressing the drawback of [44], proposed SIMPATH for influence maximization under the linear threshold model by incorporating several optimizations. SIMPATH’s seed set quality is based on its spread of influence which improves the quality of seed selection significantly.

4.6.3. Models with time constraint. If we just consider the influence whether one user can make his/her friends buy one items, we can directly consider the success of the later transaction. However, if the influence does not reflect directly, how to measure the influence of them? Consider one instance on Facebook, after user Mike posted a new status “I got a new Kindle Fire HD from Amazon, it is awesome!” with picture, besides users specific block his news feed, all Mike’s friends and followers will get this information from their Facebook Timeline and related search result. Obviously, not all neighbors who have been influenced will forward Mike’ post, but they might have already be influenced by this status. For each event, whether the action of the action from the original user can influence others or not depends on different situations. The phenomena of time-delay in information diffusion has been explored in statistics. Observation by Moro et al. [116] showed that the heterogeneity of human activities controlled the dynamics of information diffusion. Several works have been proposed to deal with the time issues. However, since the influence itself is a dynamic process, time is hard to capture. Based on the IC model, Chen et al. [41] extend the influence maximization problem to have a deadline constraint which can partly reflect the time-critical effect. IC-M is their model to capture the delay of information propagation in time which is easy to develop a (1-1/e)-approximation algorithm to circumvent the NP-hardness, and similar techniques from their previous works [43, 42] were used to compute the influence in arborescences structure. However, only a probability weight to simulate the time delay is hard to be persuasive, and the probability is just come from a random number which does not conform to the actual. Saito et al. [205] extend IC and LT to incorporate asynchronous time delay and investigate. Two models called AsIC and AsLT are proposed. Different from the work of Myers et al. [181] which focused inferring the structure of network, Saito’s approach can effectively learn the model parameters from a limited number of observed data. Thang et al. [58] model the influence maximization by limiting the influence of nodes that are
Algorithm VirAds was proposed which guarantees a relative error bound of $O(1)$ if the network is power-law, and they also provided the theoretical analysis to show how hard the model they extend to obtain a near optimal solution within a ratio better than $O(\log n)$. With an emphasis on the time efficiency issue, Chen, Zhu et al., [45] developed a framework CIM to tackle the influence maximization problem by community-based techniques. By exploiting the properties of the community structures, CIM is able to avoid overlapped information and thus efficiently select the number of seeds to maximize information spreads. Based on the continuous time model introduced in [196], Rodriguez et al. [198] improve their work accounts for temporally interactions in a diffusion network which allows information to spread at different rates across different edges.

4.6.4. Models with competition. In [37], Chen et al. introduced the extension problem of IC as IC-N which consider the negative opinions appearing in the social networks. Their model incorporates the negativity bias which means that the negative opinions usually dominate over positive opinions. Similar as the classical IC model, their model still keep the submodularity which allows the greedy approximation algorithm keep the ratio of $1 - 1/e$. They utilized a tree structures to develop an efficient algorithm to compute the influence. Nam et al. [188] studied $\beta \frac{1}{n}$-Node Protector problems which aim to find the smallest size nodes set whose decontamination with “good” information provides at least $\beta$ disinfection ration on the whole network. While different from [37], the good information has a stronger power to influence. When positive and negative information appeared at the same time, the good one will win.

The models above only consider two different opinions which incorporate negative relationships. Thus this is a simple version of competitive. Generally, since the competitors might be more than two in practice, many literatures consider more competitors in the influence maximization. Bharathi et al. [18] extend their past work by focusing on the case when multiple innovations are competing within a social network.

Kostka et al. [129] examine the diffusion of competing rumors in social networks. Game theory and location theory are used to provide the rumors diffusion process as a strategic game. Under a game-theoretic framework, they show that finding the optimal strategy of both the first and second player is NP-Complete problem.

Barbieri et al. [16] extended both IC and LT to topic-aware models which result to be more accurate in describing real-world cascades than traditional ones. A topic-aware Independent Cascade model (TIC) is proposed with the proceeds that when a node $u$ first clicks an advertisement $i$, it has one chance of influencing each inactive neighbor $v$, independently of the history thus far. And the probability of success influence is the weighted average of the arc probability with regard to the topic distribution of the advertisement $i$. Aslay, Lu, et al. [16] extended the work of Barbieri et al. [16] with Click-Through Probabilities (CTPs) for seeds. Taking advantage of network effect and paying attention to some piratical factors such as relevance of advertisement, effect of social proof, et al. [10] introduce a problem domain of allocating users to advertisers for promoting advertisement posts.

Most recently, Datta et al. [55] proposed a axiomatic approach based on cooperative game theory to define the influence measure. Their approach take the advantage of the algorithm’s independence of the underlying structure for classification function. Based on the theoretical result of this technique, experiments show
that their framework could identify advertisements where certain user features have a significant influence on whether the ad is shown to users or not.

Table 1: Extensions or improvements of IC/LT models

| References      | Extender                                                                 | IC | LT | Remarks                                                                 |
|-----------------|--------------------------------------------------------------------------|----|----|-------------------------------------------------------------------------|
| Goyal [71]      | Learnt probability from the action log, simulation on both IC and LT models | √  | √  |                                                                         |
| Chen et al. [42]| Address the scalability issue, they proposed efficiency heuristic algorithm by restricting computations on the local influence regions of nodes | √  | ×  | Showed that computing influence spread in the independent cascade model is #P-hard problem |
| Chen et al. [41]| Extended the classical IC model to study time-delayed influence diffusion | √  | ×  | Their technical report version paper provides the NP-complete hardness of LT with their time-delay feature |
| Masahiro et al. [127] | Improved the basic IC and LT by estimating marginal influence degrees | √  | √  |                                                                         |
| Chen et al. [43] | Degree discount heuristics achieve almost matching influence thread with the greedy algorithm, and run only in milliseconds which the traditional method run in hours | √  | √  |                                                                         |
| Wang et al. [236] | Heuristic algorithm for IC model                                         | √  | ×  |                                                                         |
| Chen et al. [37] | Extended the classical IC model to incorporating negative opinions       | √  | ×  |                                                                         |
| Nam et al. [188] | Focused on how to limit viral propagation of misinformation in OSNs       | √  | √  |                                                                         |
| Wang et al. [239] | Extended IC to mobile social networks, and use a dynamic programming algorithm to select communities then find influential nodes | √  | √  |                                                                         |
| Kyomin et al. [121] | Algorithm IRIE where IR for influence ranking, and IE for influence maximization are proposed to improve the classical algorithm developed previously | √  | ×  | The algorithm was used in both classical IC model and the extension IC-N [37] |
| Thang [58]      | Extended the LT model by constrain the influence distance as constant d  | ×  | √  |                                                                         |
4.7. Heterogeneous models. From different views, as shown in 10, literatures have paid a lot attentions to influence maximization in heterogeneous networks [238, 266, 157, 254, 16]. This kinds of algorithms are based on the observations that (a) users in network might have different interests, (b) the topics or items have different characteristics and (c) similar users (items) are interest to same items (users).

In the Heterogeneous, Tang et al. [220] try to learn the influence probabilities from the structure and the similarity between nodes in the social networks. They proposed Topical Affinity Propagation (TAP) to model the topic-level social influence on large networks. Additionally, their (TAP) is designed with efficient distributed learning algorithm which is implemented and tested under the Map-Reduce framework. By taking into consideration topics, Chen, Fan, et al. [36] propose a
sample-based algorithm with $\epsilon \in (0, 1]$ to maximize the influence. Another sampling diffusion model are proposed by Yang, Tang, et al.\cite{251}. They develop an active learning technique to alleviate the problem for how to collect sufficient labeled samples for training an accurate classification.

To address the problem of mining the strength of direct and indirect influence between nodes in heterogeneous networks, Liu et al. \cite{157} proposed a generative graphical model which utilized the heterogeneous link information and the textual content associated with each node in the network. Based on the learned direct influence, Liu et al. also studied the influence propagation and aggregation mechanisms in \cite{158}.

From social psychology to a computational field, a Role-Conformity Model (RCM) is proposed by Zhang, Tang, et al. \cite{261} to model the conformity between users by incorporating the utility function. By applying RCM on several academic networks, many evidences show the existence of correlations between people’s latent roles and their conformity tendency. And from their observation, they show that people with higher degree and lower clustering coefficient are more likely to conform to others. And this result could be one explain of the phenomenon that collaborations between the neighbors in the local network are infrequent.

In \cite{38}, the authors pointed out that most of influence maximization research only utilize an individual’s ability to influence another but ignores individuals’ conformity which is a person’s inclination to be influenced. Two models $C^2$ and $C^3$ are proposed to support their observation. Similarly, by adopting a linear and tractable approach to describe the influence propagation, Liu, Xiang, et al. \cite{159} developed a “Group-PageRank” metric to quickly estimate the upper bound of the social influence.

A class of diversity measures to quantify the diversity of influenced crowd are proposed by Tang, Liu et al. \cite{219}. In this work, a simple greedy algorithm with a near-optimal solution are provided to answer the question that in real social network who is influenced and how diverse the influenced population is. Considering the similarity and influence in heterogeneous networks, Wang et al. \cite{238} introduced a framework which computes social influence for one type of nodes and simultaneously measures the similarity of the other type. Similarity score and influence score are used to measure the similarity and influence score more precisely. Similarly, considering the similarity, Zhou and Liu \cite{266} introduced a vertex similarity metric in terms of both self-influence similarity and co-influence similarity. With a dynamically refine cluster algorithm, they continuously quantified and adjusted the weights on self-influence similarity and on multiple co-influence similarity scores towards the clustering convergence.

Another topic aware mode proposed by Li, Ding et al. \cite{144}. By integrating both topic factor and opinion influence factor into a unified probabilistic framework, they build a topic-level opinion influence model (TOIM). From a new perspective of sentiment analysis, they capture user opinion on different topics in heterogeneous social networks. As more and more social data are available from social media, the influence analysis is not limited to the basic relationship between users or groups but also evolve more semantic of media content themselves.

Most of heterogeneous models for influence maximization only consider the topics or users’ role in social network. Li, Chen et al. \cite{145} proposed one location aware model to maximize the influence. As the development of mobil applications, location is not a unobtainable information any more. Many real-world applications such as location-aware word-of-mouth marketing also have location-aware requirement too.
To solve the influence maximization problem in a heterogeneous information network which combing the data from both sensed cyber-physical world and online social world, the comprehensive resolutions are proposed in [88, 87]. Four behavior patterns and corresponding formulated functions are proposed to model the users’ behavior in sensed cyber-physical world [166]. By adopting the classical influence maximization technique and differential privacy, the approaches can achieve an efficient influence maximization algorithm with privacy protection. The real life data experiments verified that the framework works well for the problem of influence maximization and the proposed algorithm is outperformed other up-to-date resolutions.

4.8. Epidemic model. The spread of disease has been studied for many years by biologists. Similar to the disease, the information spread follow the process of one suspecting another, and passing on. The social influence model based on biological transmission is shown in Figure 11. Newman and Mark proposed so-called susceptible/infective/recovered (SIR) model [185], which describe the spread of a disease on network. In (SIR), individual occupies one of the three states, “susceptible”, “infective”, and “recovered”, where a susceptible individual becomes infected with a probability when an infected patient and subsequently recovers at a rate. Similar techniques have been noted from the computational biology [167]. We can easily find that the classical IC model can be identified with (SIR), where the nodes become active at time $t$ in IC model correspond to the infective nodes at time $t$ in (SIR). Therefore, the IC is equivalent to a percolation model, and probability distribution for the final active nodes in these two models are same. The techniques of computing the influence (SIR) have also been used to improve classical IC [127][81].

Different from the (SIR) model, the SIS model where the last “S” denotes susceptible again actually is a general version of SIS. In (SIR) model, only infected individuals can infect susceptible individuals, while recovered individuals lost the probability of infecting others, and they can not be infected by others also. As Saito et al. [204] pointed out that more applications such as the growth of hyperlink posts among bloggers [141], epidemic disease spread [186] and the prorogation of computer viruses which can be more appropriately to use the SIS model. Through a Markov process, [7] introduced a analytical information dissemination model which extend the epidemiological model SIS. Cannarella, John et al. [28] developed epidemiological model named ir-SIR to tackle the dynamics of network. They modify the traditional SIR model of disease spread by incorporating infectious recovery
dynamics such that connection between an infected and recovered user of the network is required for recovery. They case study of Google search query for both “MySpace” and “Facebook” both exhibited the abandonment phase of their model.

As shown in Fig. 11, more models related to disease spread can be considered[230]. Li, Bhowmick, et al. [146] recently specialize the fact that most of influence maximization model that only utilize an individual’s ability to influence another but ignores individual’s conformity which is a person’s inclination to be influenced. Based on the model they proposed, they could provide a feature that influence are aligned to the popular social forces principle in social psychology.

A graphical game model is introduced by Huang [114] to analyze the information diffusion system. Nash equilibrium has been invested to get more new discovery such as the user with higher valuation are welling to make more effort to enrich the original information.

[1] address the reconstruction problem formally. A union of several networks individually generated from metrics is structurally different from networks generated from just one metric. They provide a near-linear algorithm for reconstructing the latent social structure with provably low distortion. The model explicitly produces a union of graphs with one graph for each category. An important feature of the algorithm is that it separates the different graph from each other. The result of their work can be interpreted as a proof of concept that it is possible in principle to efficiently separate the different dimensions of social interactions and identify similarities between individuals.

Traditionally, it was hard to capture and study the effects of mass media and social networks simultaneously [155]. However, the Web, blogs, and social media changed the traditional picture of the dichotomy between the local effects carried by the links of social networks and the global influence from the mass media. In [138], the authors develop a framework for tracking short, distinctive phrases which has potential relationship among online text.

[143] considered the situation that when one selected influential seed has been removed and what is the the best strategy to select the successor node to replace the removed one. Zhang, Chen et al. [263] consider the task of selecting initial seed users of a topic with minimum size to approach the number of users discussing the topic would reach a given threshold with a guaranteed probability.

Recently, to extend both topic-aware and efficiency issue, Li, Zhang et al., [152] propose a keyword based targeted model for online targeted advertising. The model try to find a seed set that maximize the expected in fluency over users who are most relevant to a given advertisement.

Yang, Tang et al. [250] studied the interplay between users’ social roles and their influence on information diffusion. As another kind of heterogeneous feature, social roles is one of the most important features in social network. Moreover, social roles are not independent of information diffusion in nature. One sampling based algorithm are developed in this paper to learn the proposed model using historical diffusion data. After the verification of an experiment on real data from Tencent Weibo, they expect that their model could be applied to different scenarios to predict the scale and the duration of a diffusion process.

4.9. **Theoretical result of influence maximization.** Since Kempe et al.[124] formulated the classical models $IC$ and $LT$, then they gave the proof of the NP-hardness and provided an approximation algorithm for selection of influential nodes.
Based on the result of Nemhauser et al. [184], a monotone and submodular function $\delta(\cdot)$ can obtain an approximate greedy algorithm with factor of $1 - 1/e$. Mossel et al. [180] generalized the results of Kempe et al. [124] then in [179], with Roch, Mossel give a better result which improve the theoretical result of approximate algorithm ratio for influence maximization from $(1 - 1/e)$ to $(1 - 1/e - \epsilon)$ where $\epsilon \geq 0$. They also state that in the classical model, when influence between individuals is submodular, the same to the objective function in the global influence maximization.

After reaffirming the result of [179], a fractional version of the influence maximization problem has been proposed by Demaine et al. [57]. Different from the binary choice (stoats include active and inactive) of the classical model, the users can been partially influenced where the classical can be seen as a special case of it. Similar idea is another version of this kinds of extension which times the influence but not segments the influence.

Based on one phenomenon in real world that one user could accept or buy the same item multiple times, Lu, Wei, et al. [164] proposed a propagation model MIMA for influence maximization. Different from other traditional models, MIMA consider more multiple actions in real world. One conception acceptance volume is introduced in their model, and they are aiming to maximize the overall acceptance volume of all the activated nodes for an item based on their model. Another work concern the repeat influence activation called cumulative influence maximization was proposed is proposed by Zhou, Zhang et al. [265]. Different from [164], [265] does not use multiple acceptance volume but cumulative influence to measure the influence propagation, then find out the best initial seed set to maximize the influence.

One of the most recently theoretical result of influence maximization is proposed by Christian et al. [21], which is an $O\left(\frac{(m+n)\log n}{\epsilon^3}\right)$ algorithm with approximation ratio $1 - \frac{\epsilon}{e} - \epsilon$. However, Chen et al. showed that both computing influence spread in the independent cascade model [42] and linear threshold model[44] are $\#P$-hard problems. Even greedy algorithms cannot be finished in an acceptable time. The heuristic algorithm and strategies will be surveyed in the next subsection.

Sanjeev and Brendan [125], similarly, consider the problem of finding $k$-size maximum influence nodes in the undirected network. They extend the result of traditional IC model to undirect network, and achieve an $(1 - 1/e + c)$ approximation to the set of optimal influence for some $c > 0$. In their theoretical analysis part, they also show the APX-hard of the influence maximization problem.

All works above consider more or less the submodularity and monotonicity of the influence maximization. Besides the heuristic algorithm, they apply the hill-climbing search to achieve the approximation ratio. Most recently, Zhang et al. pointed out a variant influence maximization problem, and give the theoretical proof that the problem does not follow the same submodularity when the objective function goes to a probabilistic coverage guarantee.

Two sampling models are proposed by Tang et al. [223] to sample the representative users. They give a formal definition of the problem and try to find a subset of users to statistically represent the original social networks. Their experiment shows that it only take a few seconds to sample 300 representative users from a network of 100,000 users. However, the construct the representation of users depend on the specific attribute such as numeric attributes and non-comparable attributes.
Their method could be applied to some semantic networks with background, but not represent the structure features of network.

Another attempt to tackle the hardness of the possible world is from Panos, Francesco et al. [193]. Their method aim at preserving the expected vertex degrees since these feature capture the graph topology well in practice. After applying conventional processing techniques on these representative instances, their method could closely approximate the result on the uncertain graph. Recently, He and Kepme [103] proposed a new paper which prove the submodularity of influence difference maximization for the IC and LT models, which is one omnibus result to their classical models IC and LT.

Borgs et al.’s [21] method first avoid the limitation of traditional greedy algorithm. Their research shows a drastically different technique for influence maximization under the IC model. From the perspective of the opposite, [21] define a reverse reachable (RR) set for node $v$ in the network is the set of nodes that can reach $v$. Then by sampling algorithm, the algorithm generated a certain number of random possible world of (RR) sets from the network. Follow the rationales that if a size-$k$ node set $S$ could covers most (RR) sets, then $S$ has a higher probability to maximize the expected spread among all size-$k$ set in the network. Their theoretical result shows that when parameter $\tau$ is set to $\Theta(k(m+n)\log n/e^3)$, the algorithm could run in time linear to $\tau$, and returns a $(1-1/e-\epsilon)$-approximate with a constant probability. Further more, Tang, Xiao et al., [226] proposed a more practical framework TIM which guarantees the same theoretical complexity bound and keep at least probability $1-n^{-l}$. TIM supports a triggering model, which is a more general model includes both IC and LT as special cases.

4.10. **Heuristic strategies for influence maximization.** Heuristic strategies could provide a very efficient algorithm with very cheap computing cost. Some straightforward strategies are amiable for influence maximization. Although most of them have their defects of natural such as unmeasurable or poor universality, the idea of heuristic method is still very important to adapt to some specific situation.

4.10.1. **Random strategy.** As a baseline comparison of most algorithms in influence maximization, randomly selecting $k$ vertices in the network can been considered as the simplest strategy. Although this idea is very simple, it can also give a uniformity result if the relationships in the network are relatively balance. Additionally together with the ease of implementation, this method are also used and compared in many literatures.

4.10.2. **Degree priority strategy.** This strategy greedily selects the highest degree node to the potential seed set until the process meets the stopping condition. This is one of the most naive method to find the most influential nodes, but in some groups or communities, the average degree might be much higher than other part. To maximize the influence nodes set in a network with a size constrain by this strategy have to be limited by this feature.

4.10.3. **Degree discount priority strategy.** As an extension version of the basic degree priority strategy, this heuristic algorithm choose the largest degree node $v$ each step, and after adding $v$ to the active seed set, all the neighbors’ degree of node $v$ will be reduced by 1. This is one plus version strategy of basic degree priority, but both degree based strategies do not take the strength of nodes’ relationship into account.
And since the influence diffusion is a process rather than one step deal, one high degree nodes could not result in a large area influence.

4.10.4. Shortest path based strategy. Shortest path is another strategy for influence maximization. The main idea is to find nodes in the network that can reach as more other nodes as possible with a shortest path. At this point, the shortest path represents the influence route of nodes. If one node could reach a lot of other node with a very short path, it indicates that this nodes should have a better influence. One problem of this strategy is that one node could have many different paths to reach another node. Even with the same length, it is very hard to control the strategy when we consider more than one pair of nodes.

4.10.5. PageRank. PageRank [191] is one of the most foundational research results of Google’s founders Page and Brin which bring the order to the web. The basic idea is to sort the web pages by readers interests, knowledge and attitudes. For influence maximization, the PageRank strategy could also be employed to rank the nodes in the network, and give each node a score similar to PageRank but based on the activity of node, then output the nodes with the highest score. The weakness of PageRank is that this method only gives score to individual node rather than a nodes set. Therfore, a certain high score node could not result in the final high influence. Moreover, some high score nodes might cluster together which also reduce the entire influence of the result set.

Recently, another ranking-based strategy $IMRank$ is proposed by Cheng, Shen et al. [46]. $IMRank$ finds a self-consistent ranking by reordering nodes iteratively in terms of their ranking-based marginal influence spread computed according to current ranking. A last-to-first allocating strategy are proposed to improve the efficiency of estimating the marginal influence for a given ranking. By adopting a linear and tractable approach to describe the influence propagation, Liu, Xiang, et al. [159] develop a “Group-PageRank” metric to quickly estimate the upper bound of the social influence.

From a different point of view, [175] considered the situation that for real application, obtaining complete knowledge of a social network’s topological structure is not a easy work, thus, the authors take the problem of $IM$ in unknown graphs, and propose a heuristic algorithm for the problem. In their problem, the social network’s topological structure is initially unknown, only the number of nodes is given, and a limited amount of probing is allowed to obtain a partial structure of the social network. Biased sampling strategy (snowball sampling strategy) is applied to probing the network.

There are also many other heuristic techniques applied to some classical algorithms as part of supplement to improve their efference or reduce the computation complexity. As one important and useful member in family of influence maximization algorithms, to develop a effectively and efficient heuristic algorithm is always one good alternative choice for researchers. As shown in Figure 12, the comprehensive model structure of influence is demonstrated.

5. Applications motivated by influence analysis. Multiple applications are motivated by influence analysis, as shown in Figure 13.

5.1. Influence analysis in mobile social networks. Han et al. [83] studied the target-set selection problem for information delivery which serves for the bootstrapping mobile data offloading. They propose a heuristic algorithm to select a
target set with size $k$ among all subscribed users, thus maximizing the number of users that receive the delivered information through the mobile opportunistic communications [27, 26]. Recently, they also study how to identify the influential users in mobile social networks [84]. Different from other methods, they propose a distributed protocol through fixed-length random walks which can be used on smartphones and identify influential mobile users. Nguyen et al. [187] present a framework to adaptively update the community structure by selecting critical nodes in dynamic networks [105, 106].

Wang et al. [239] proposed an algorithm called Community based Greedy algorithm for mining top-K influential nodes in a mobile social networks. They extended the basic $IC$ model to take weight edge into consideration. By taking information diffusion into account, their algorithm first detect communities through dividing the social network into smaller communities. Then find influential nodes from selected communities by a dynamic programming algorithm.

Song, Zhou et al. [215] proposed a divide-and-conquer method to do the influence maximization on a large-scale mobile social network. Parallelized computation mechanism has also been adopted in their method to tackle the efficiency on their large-scale mobile network which has 26 million edges and around 5 million nodes.
Yang, Jia, et al. [249] consider the emotion disclose problem from image in social network. Different from other emotion analysis method, the emotion analysis in their paper is a learning based method by jointly modeling images posted by social users and comments added by their friends. And this model could distinguish those comments that are closely related to the emotion expression for an image from the other irrelevant ones.

[85] presents two novel models TIC and TLT which extend the practicality of the classical IC and LT models for influence maximization. The theoretical analysis shows that the two new models they propose both follow the monotonicity and submodularity. This result could help us to design simple greedy algorithm with a guaranteed approximate ratio $(1 - 1/e)$. Both the synthetic and real social network data are tested by the implementation on Hadoop and Spark platforms, showing that the algorithm for TIC and TLT could solve the problem efficiently and effectively.

5.2. Influence analysis for emotion prediction. Tang et al. proposed their approach for the emotion prediction problem which aims to study individual’s emotional states evolve in social network systematically and quantitatively [224]. By using a data set including 36,000 hours of continuous behaviors and emotional states from the mobile phones of 30 users, they observed that the influence of different time was generally based on the previous time, and the user’s emotional state might also be influenced by their friends. Recently, Xia et al. [111] make the sentiment analysis in Microblogging which investigated the social relations and considered the influence in the social network.

5.3. Influence analysis for recommendation. A series of online experiments have been developed to investigate whether online recommendations can sway user’s opinions. Their results show that people’s own choices are significantly influenced by the perceived ideas from others. However, the effect is weaker when people have just made their own choices. Additionally, the first decision user has made significantly predict whether they will reverse their own opinions later on [267].

Considering the patent partner recommendation in enterprise social network, Sen et al. [244] proposed a framework in an online model which incorporate users’ interactions. By the framework they proposed, they try to figure out what are the fundamental factors that influence the co-invention relationships. Focusing on the case that the seed users who are targeted such as the new product and endorse it with relatively high ratings, a novel problem RECMAX has been proposed from Goyal et al. [75]. RECMAX aims to find a set of seed users to offer them a earlier promotion then let the recommendation from them to maximize the market.

Besides aiming to recommending connections by the number of common neighbors and similarity of user profiles, etc., authors of [34] proposed algorithms to boost content propagation in a social network without compromising on the relevance of the recommendations. Instead of nodes, they were looking for edges with a bound on the number of incident edges per node. They also proved that the content spread function is not submodular, and proposed approximation solution for computing the near-optimal set of edges. The authors of [126] identified the impact of social influence in various aspect of E-commerce and introduced how to exercise social influence on customer’s decisions. Ida et al. [174] presented a graph-based data abstraction for modeling the user behavior through browsing. They focused on news and blog pages, which are more appropriate for recommendation. Although extensive studies
have been paid for addressing the prior expectation recommendation, less attention has been focused on investigating the users’ posterior evaluation. The authors of [115] find a counter-intuitive phenomenon that word-of-mouth recommendations are strongly related to users’ posterior evaluation. They proposed a framework to quantitatively measure individual’s social influence by evaluating the number of users’ followers and their sensitivity of discovering items, and further verified that the raise of the posterior evaluation is directly caused by word-of-mouth recommendations. In heterogeneous, Xiao et al. studied building hybrid recommender system by using additional user or item relationship information[254]. With user’s feedback, they propose to combine various relationship from the network together. As a friend recommendations system, Yang et al. [248] proposed the Acceptance Probability Maximization (APM) problem, which is also based on the influence and interaction analysis in social network.

Another kind of recommendation is named as “cold-start” [135], which means that less or no history knowledge we could learn to do the recommendation. Therefore, it is hard to analyze the influence in the network. This situation has a similar feature as the beginning of some in-time influence diffusion model in big dynamic environment. One potential solution is proposed by Zhang, Tang, et al. [262]. They address the cold start recommendation with a semi-supervised co-training algorithm which also provides a flexible way to incorporate the unlabeled data.

Another result for cold start recommendation proposed by Rong, Wen et al. [200] takes a precomputation approach, and computes the user’s similarity to predict the rating for the new users. All these kind of recommendation analysis the influence between users in the social context.

5.4. Influence analysis for communities. Communities is one of the most important feature and nature properties in real social network. But how to quantify node’s local influence is always a challenging question. Jiang, Jin et al. [119], based on influence maximization, a powerful tool to detect communities, develop a uniform framework for community detection in social network. Their techniques employ local influence maximization as the community formation process, then use local influence as a measure of evaluating node importance in its local neighbors to detect all communities.

Most of previous works on influence maximization modeling with topic-aware have assumed one-to-one correspondence between communities and topics. But since rich correlation between communities and topics are ignored, it limits the practical utility. Most recently, Hu, Yao, et al. [112] proposed COLD, which models topics and communities in a unified latent framework. COLD uncovered and explored temporal diffusion and extract inter-community influence dynamics. In addition to this, by associating each community with a mixture of topics, COLD can explore communities’ varying topical interest. Li, Qin, et al. [149] considered the influence of a community in a network and addressed the problem of finding densely connected subgraphs that satisfy the query conditions. However, their method is based on the concept of k-core and network structure model, which is not based on the influence diffusion model. This kind of model could not provide influence expectation measures, thus could not completely follow the information diffusion process as proved in many other literatures.

Taking a conventional social network activity as an example to discuss influence diffusion in daily life. Assume there is one user on Facebook sharing a new song or movie. This action results in an influence diffusion process. That is, friends or
followers of the action initiator will have similar behaviors - be influenced. Consider one instance as an example. User Mike posts a new status “I got a new iPhone 7 plus from Apple Store with student promotion. It is awesome!” with pictures on Facebook. All of Mike’s friends and followers will get this information from their Facebook’s news feed or related search results. The effect of this post will be weakened as time goes on. For acceptance ratio, obviously not all the neighbors who see the post will forward it. Although some of Mike’s friends might have already been influenced and begun to take next step to purchase an iPhone, some of his friends might have simply ignored this post. Considering the receiving of that post as the first step of influence, all the users having a friend relationship with Mike have a possibility to receive this influence. But only the neighbors who comment, forward this status, or take response action regarding this post could be considered as accepting the influence, which is the second step of the influence. For the breadth of influence, one possibility is that a lot of Mike’s friends are studying at the same department of the same university. If we evaluate the influence ability of Mike in the whole social network, he might not be as good as another user Michael, who has fewer friends studying in many different universities. Compared with Mike, Michael has a good chance to pass the influence much more broadly than Mike. Consider the coverage of influence diffusion, a practical probing framework to explore the dynamic of networks in [94]. The probing framework takes the community as a unit and updates network topology by only probing $b$ communities instead of searching the entire network. Besides, a divide-and-conquer strategy is applied with dynamic programming technique to maximize the community-based influence. The comprehensive experiment results show that the model can achieve comparable influence diffusion performance compared to the node-based probing algorithm while having much better efficiency and more applicable to large-scale networks. Specifically, in the extended version of their model, the authors use the number of communities to measure the breadth of the influence, which is novel.

5.5. Influence analysis with parallel techniques. A framework is proposed to accelerate the influence maximization by leveraging the parallel processing capability of graphics processing unit (GPU). In their work, a bottom up traversal algorithm was proposed to improved the basic greedy algorithm by converting the graph into a directed acyclic graph to avoid deadlock and calculate influence spread based on their child nodes. An adaptive $K$-level combination method was further developed to maximize the parallelism and reorganize the influence graph to minimize the potential divergence [160]. Considering an independent influence path as an influence evaluation unit, an approximation algorithm named as Independent Path Algorithm (IPA) were proposed to approximates influence. The parallel versions of their IPA speeds up further as the number of CPU cores increases, which can been adapted to a larger size of datasets [253].

Tang et al. also made many efforts to accelerate the efficiency by paralleling the algorithms. COLD [112], which is the model focusing on the community level influence diffusion extraction, provides a parallel inference implementation on GraphLab. Regarding to a large scale data, [169] considered the task of evaluating the spread of influence in large networks in IC model and studied the question of designing scalable algorithms for estimating cascades under the same model. The main idea of [169] is to estimate influence via a sampling approach that allows both parallelization and trading off between simulation cost and informativeness. And the
probabilistic analysis is also employed to illustrate how a algorithm can choose parameters to navigate this tradeoff appropriately.

5.6. Applications for other specific networks. Twitter, as one of the most famous micro-blog web site, many algorithms are developed to analyze the data of it [227, 13, 64, 203]. Based on the evaluation on Twitter, Romero et al. [199] claimed that making individuals to become influential not only need obtain attention and be popular, but also be necessary to overcome user passivity. Weng et al. [241] developed a topic-sensitive PageRank named TwitterRank which taking the topical similarity between both the user and the link structure into account to measure influence in Twitter. They claimed that their study reveals the presence of “reciprocity” that can be explained by the phenomenon of homophily [173]. Cha et al. [30] performed a comparison of indegree, retweets, and user mentions as three different measures of influence on Twitter. Different from the work of Weng et al. [241], their result show that the reciprocity is low overall in Twitter. They also investigated the dynamics of user influence across topics and time. Based on the observation, they argued that influence is not gained spontaneously or accidentally, but through concerted effort. Users in social network need to keep great personal involvement to gain and maintain influential. The authors of [128] studied the retweeting convention adopted by more than 2 million people in the popular social network Twitter, they got the similar result that the practical action are influenced by their friends. Considering the indirect influence in Twitter, Xin et al. [211] proposed a quantum cognition based probabilistic model to account for local drops which come from their observation. They also investigated the propagation of parallel indirect influence on Twitter with considers the number of spreaders. A Twitter context tree is build by Chang et al. [33] to help users understand the contextual information. They studied how to improve summarization methods by leveraging the rich user interactions and they proposed a Granger Causality Influence Model to model the time series influence in Twitter. Considering discriminative influence, UDI, a unified discriminative influence model, were proposed by Rui et al. [150] to profiling users’ home locations in Twitter. Jing et al. [259], in another point of view, studied the phenomenon of social influence locality in Twitter. Based on pairwise influence and structural diversity, they provide two instantiation functions which help to understand the underlying mechanism of users’ retweet behavior influence with each other.

Besides the data of Twitter, data of Youtube [252], Flickr [177][32], and Facebook [14] etc. were also considered by literatures. The authors of [252] examine how the size and structure of the local network around a node affects the diffusion of products seeded by it in the context of YouTube.

Tao et al. [218] focused on the online discussion forums and proposed the participation maximization problem, which is another specific influence maximization. Although approximation and heuristic algorithms were developed, they still faced the NP-hard challenge. Different from IC, based on the influence, their model is for user appending posts to existing threads. Meeyoung et al. collected and analyzed large-scale traces of information dissemination from Flickr which is one of biggest photo share platform [31]. Mislove also analyzed the Flicker from the network growing view [177]. Scholars from Facebook [14], examine the role of social networks in online information diffusion with a large-scale field experiment which among 253 million subjects. They further examine the relative role of strong and weak ties
in information diffusion and pointed out that even stronger ties are more influ-
ential individually, it is the more abundant weak ties when the novel information
propagation.

Focusing on three different graph classes: Erdős-Rényi, planted partition and
gеometrically structured graphs, Ok, Jin et al. [189] propose polynomial time ap-
proximation algorithms with a guaranteed approximated ratio in $O(n^2)$ time. They
follow a game-based diffusion model which motivated by the observation that peo-
ple’s behavior is often strategic when they decide to adopt or not the innovation (i.e.,
individual follow the innovation only if it provides sufficient utility, which changing
with the choices the neighbors adopting operation).

Considering the network which extracted from MEDLINE, a network-based al-
gorithm which ranks heterogeneous objects is proposed by [35]. They try to figure
out the most influential literatures from the MEDLINE. Other literature consider
the famous co-author academic publications network such as DBLP6, Arnetminer7
also provide many influence diffusion and influence research results.

From a very interesting aspect, Dong, Johnson et al. [61] analyzed the scientific
impact of citations in their research by considering the measure $h$-index, and tried
to answer the question “Will this paper increase my $h$-index?”. Two factors, the
authors’ authority on the publication and publication venue, play the most decisive
roles and are proposed to contribute to the primary author’s $h$-index. However,
the popularity of publication topic and co-authors’ influence are surprisingly not
strongly correlated to the prediction target.

5.7. Other applications. Tang et al. [222] analyzed a special type of social influence
which involves a change in opinion or behavior in order to fit in with a group
called conformity. A model Confluence was proposed to formalize the effects of
social conformity into a probabilistic model. Effects of the different types of con-
formities can be distinguished and quantified by their model. To scale up to large
scale networks, they also proposed a distributed learning method to speedup their
Confluence model. Another group influence application is proposed by Feng, Kaiyu
et al. [67]. They try to identify the event organizers in online social networks.
The event organizers with special features are actually influencers in the traditional
question [171].

In [40], Chen et al. used the well-known Bayesian Nash equilibrium, tried to
maximize the selling of a digital product in a social network by choosing the price.
Considering the situation that the information of the network is incomplete, sam-
ping techniques are applied in their algorithm. Yaron [212] introduced mechanisms
that elicit individual’s costs while providing desirable approximation guarantees in
some of most classical models of influence. His target is not just following the ba-
sic influence maximization model, but also winning friends and influencing others
in a truthful way. Bhagat et al. [17] adapted the classical $LT$ model by defining
an objective function which captures product adoption. The model they proposed
still keep the monotone and submodular. Further, an approximation algorithm was
introduced to solve their application. To maximize the adoption of a new product,
Barbieri and Bonchi [15] study the problem of designing the features of a novel
product. Based on influence maximization, they model different products to different
characteristics, then maximize the adoption of product.[101] Jiang, Jin et al.

6http://dblp.uni-trier.de/db/
7http://arnetminer.org/
[119], based on influence maximization, develop a uniform framework for community detection in social network.

6. **Future research directions.** There are several exciting directions to pursue around the influence analysis in social networks. In this section, we are going to show more future problems and challenges for influence analysis, and further point out the future research directions in the following.

6.1. **Competitive influence analysis.** All topics we discussed are working on single information resource, but there are many different situations in real life that more than one information resources are existed.

Bilateral competition diffusion model could be considered as the two opposite opinions in the social scenario, where one is positive while the other is negative. How to analyze the information diffusion is a very challenging and meaningful research topic. In real life, it is quite common in the situation that different ideas are competing for their influence in the social networks. Such competing diffusion could range from two competing companies, friend and foe relations, two political candidates of the opposing parties to even the government tries to inject truth information to fight with rumors spread to the public. From a competitive aspect, besides the bilateral competition models, the competitors could be more than two. For example, BMW, Ford, Honda, Toyota, and Tesla are all famous car brand. How to model the influence of multiple competitors are very challenging. These scenarios are arising in the real world. In many companies with comparable products, more than two political parties run for the election. How to model many different competitors with or without confliction in a social network to propagate the influence is still a very challenging problem. Map coloring and game theory might be very potential resolutions for multiple competitors influence problems. However, there is still no practical resolution available. How to solve these kinds of problem might be one of the important further research directions in influence analysis.

6.2. **Influence analysis with domain knowledge.** Domain knowledge could be used to refer to an area of human endeavour, an computer activity, or other specialized discipline. Incorporating many kinds of domain knowledge could greatly enhance the ability of influence analysis techniques.

With the development of modern mobile devices, the connection between cyber-physical network and online social network is significantly strengthened. Integrating cyber-physical knowledge into influence analysis could improve both the accuracy and practicability. However, the two kinds of data, cyber-physical and online social network, are very different from each other. How to combine the cyber-physical information and online information together to construct a novel framework for influence analysis is still an open problem. Besides the intuition problem of analysis, when we apply the cyber-physical knowledge to our problem, an imperative issue is that cyber-physical world are carrying a lot of private information of social participators.

Thus, how to analyze the influence with consideration of privacy is still a very challenging problem. For example, location information has been studied in cyber-physical aspect for a long time. Location information could significantly improve the quality of our influence analysis since if one event happens in a particular location, it could direct influence all user around that area, and this kind of influence is more specific and observable. However, the location information is very sensitive to both
users and the researchers. How to analyze the influence with privacy preserving is still blank.

Besides the domain knowledge in cyber-physical world [90, 86], the marketing knowledge could be applied to many business applications. One of most important application of influence maximization is business marketing [92]. In the domain of marketing, influence is being used to promote new product, deliver promotion, and spread marketing campaign. In political life, political views, dissent, and attitude also need to spread and expand. Influence maximization could be one of very powerful tools for political parties. In health domain, how to spread the health lifestyle and reduce the un-health habit are also very related to influence analysis.

6.3. Influence analysis in massive scale data. As the number of available data increases, kinds of massive scale data are available which offers us more and more new issues. Although a lot of challenges are standing here with us, developing new models and algorithms to solve the influence analysis problem in big data era will also be the valuable opportunity as the new big data techniques such as Hadoop, Spark, etc. appeared. From the big data, we could have unprecedented ability to figure out the influencer and the way of information dissemination which would lead to another new field of vision to see the information world. More importantly, the most notable big data platforms such as Hadoop and Spark provide us a potential solution for large scale networks to do the influence analysis. Hadoop is an Apache project and uses a distributed file system for the analysis. It provides a framework for transformation of very large data sets using the MapReduce paradigm. Hadoop is available via the Apache open source license, which provides us an opportunity to develop a big data environment for our influence analysis challenge. Spark is a very fast and general engine for big data processing. With built-in modules for streaming, SQL, machine learning and graph processing, it allows us to do the in-memory analysis for influence. How to analyze the influence in massive scale social data especially in the innovative platforms is still a very challenging question. In this case, we are going to do more research regarding the model and algorithm to investigate more potential of influence analysis. We also believe that the big data will still have great potentials and values to investigate.

New challenges appeared to both efficient and effective. For some online real-time applications, how to produce instant analysis result to the web services or to the customer is challenging. Moreover, as the data size increases, how to get even a relatively accurate result is still a challenge problem for some traditional models. Besides, one of white house reports highlighted some of the major risks in the ubiquitous use of big data technologies last year. The report mentioned that large scale data collection and analysis is glaring to be lack of transparency which need to be concerned. The security of data analysis in big data is a big topic and should not be overlooked at any time.

6.4. Influence analysis and sentiment analysis. Sentiment analysis, known as opinion mining, refers to the text analysis and computational linguistics to identify and extract subjective information in materials by using natural language processing [4, 5]. Li, Ding et al. [144] take one step to capture user opinions on social different topics in heterogeneous social network, then model the network structures, user behaviors, and user opinion preferences into a unified model to maximize the influence [221].
As more and more social data are available from social media, the influence analysis is not limited to the basic relationship between users or groups but also evolves more semantic of media content themselves. As a result, the influence maximization incorporating sentiment analysis would be another direction for further research.

6.5. Comprehensive and specific model for applications. Hopcroft et al. [109] studied the prediction problem in dynamic social network which focused on the two-way relationship. They monitored the change of the twitter network structure from 10/12/2010 to 12/23/2010. And extracted all tweets posted by the famous users they selected and in total there are 35,746,366 tweets. Based on the analysis of their data, they answer the question that “who will you follow you back?” in twitter to some extent. But the influence model they provide is limited to their data and difficult to extend. Zhang, Ariel [256] et. al., formulate a dynamic influence maximization problem to scale over a finite time horizon where a budget constraint need to be guaranteed to the decision maker. Both optimal and heuristic algorithms are proposed to solve their problem. Their model focuses on the long-term product uptake in market.

In the future, an important and challenging research area is to develop efficient, effective and quantifiable social influence mechanisms to enable various applications in social networks and social media. This area lies in the intersection of computer science, sociology, and physics. In particular, scalable and parallel data mining algorithms, scalable database and web technology have been changing the strategies sociologists use to solve this problem. Instead of building conceptual models and conducting small scale simulations and user studies, more and more people now rely on large-scale data mining algorithms to analyze social network data [120]. This provides more realistic results for large-scale applications [93, 91]. This paper provides an introduction of the problem space in social influence analysis. The area is still in its infancy, and we anticipate that more techniques will be developed for this problem in the near future.

7. Conclusions. In this paper, we defined social influence and stated its importance in evolving social networks. We introduced some analytics used when measuring centrality in social networks such as centrality measurements. We also surveyed measure models, which address the objective of influence maximization in social networks. We stated the strength and limitation of each model through a comparative study.

Social networks are graphs of individuals and their relationships, such as friendships, collaborations, or advice seeking relationships. With the increasing popularity of social networks services, more and more people communicate with each other through such networks. This survey mainly conveys a framework for studying the information diffusion problems and their approximations as well as optimizations. It provides with the readers a number of interesting models, and wise algorithms on social network.

As we have went through, novel and interesting questions thrown out by the initial work from Domingos and Richardson, inspire Kempe et al, Mossel and Roch and many others to develop a solid theoretical foundation of literature resources on the influence maximization problem. The main challenge now is to find solutions that are applicable in real viral marketing environment. Working towards various models and algorithms, researchers are trying to find a way that could really gives the satisfying result with the comprehensive experiments while without requiring too
much data load or making unrealistic independence assumptions. In order to achieve this goal and to determine the real applicability of the existing approaches, more wise designs and empirical studies are needed, and the test of the approximation techniques are also required.

REFERENCES

[1] I. Abraham, S. Chechik, D. Kempe and A. Slivkins, Low-distortion inference of latent similarities from a multiplex social network, SIAM J. Comput., 44 (2015), 617–668, arXiv:1202.0922.
[2] R. Agrawal, Nature of information, people, and relationships in digital social networks.
[3] R. Agrawal, M. Potamias and E. Terzi, Learning the nature of information in social networks, 2012.
[4] C. Ai, M. Han, J. Wang and M. Yan, An efficient social event invitation framework based on historical data of smart devices, in Social Computing and Networking (SocialCom), 2016 IEEE International Conference on, IEEE, 2016, 229–236.
[5] H. Albinali, M. Han, J. Wang, H. Gao and Y. Li, The roles of social network mavens, in The 12th International Conference on Mobile Ad-hoc and Sensor Networks (MSN 2016), 2016, 1–12.
[6] A. Anagnostopoulos, R. Kumar and M. Mahdian, Influence and correlation in social networks, in Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, Las Vegas, Nevada, USA, 2008, 7–15.
[7] C. Anagnostopoulos, S. Hadjieftymiades and E. Zervas, An analytical model for multi-epidemic information dissemination, J. Parallel Distrib. Comput., 71 (2011), 87–104, 1891295.
[8] T. C. Antonucci, K. J. Ajrouch and K. S. Birditt, The convoy model: Explaining social relations from a multidisciplinary perspective, The Gerontologist, 54 (2014), 82–92.
[9] S. E. Asch, Opinions and social pressure, Readings about the social animal, 193 (1955), 17–26.
[10] C. C. I. Aslay, W. Lu, F. Bonchi, A. Goyal and L. V. S. Lakshmanan, Viral marketing meets social advertising: Ad allocation with minimum regret, Proceedings of the VLDB Endowment VLDB Endowment Hompage Archive, 8 (2015), 814–825.
[11] D. B. Bahr, R. C. Browning, H. R. Wyatt and J. O. Hill, Exploiting social networks to mitigate the obesity epidemic, Obesity (Silver Spring), 17 (2009), 723–728.
[12] E. Bakshy, D. Eckles, R. Yan and I. Rosenn, Social influence in social advertising: Evidence from field experiments, in Proceedings of the 13th ACM Conference on Electronic Commerce, ACM, Valencia, Spain, 2012, 146–161.
[13] E. Bakshy, J. M. Hofman, W. A. Mason and D. J. Watts, Everyone’s an influencer: quantifying influence on twitter, in Proceedings of the fourth ACM international conference on Web search and data mining, ACM, Hong Kong, China, 2011, 65–74.
[14] E. Bakshy, I. Rosenn, C. Marlow and L. Adamic, The role of social networks in information diffusion, in Proceedings of the 21st International Conference on World Wide Web, ACM, Lyon, France, 2012, 519–528.
[15] N. Barbieri and F. Bonchi, Influence maximization with viral product design, Proceedings of the 2014 SIAM International Conference on Data Mining, 2014, p9.
[16] N. Barbieri, F. Bonchi and G. Manco, Topic-aware social influence propagation models, in Proceedings of the 2012 IEEE 12th International Conference on Data Mining, IEEE Computer Society, 2012, 81–90.
[17] S. Bhagat, A. Goyal and L. V. S. Lakshmanan, Maximizing product adoption in social networks, in Proceedings of the Fifth ACM International Conference on Web Search and Data Mining, ACM, Seattle, Washington, USA, 2012, 603–612.
[18] S. Bharath, D. Kempe and M. Salek, Competitive influence maximization in social networks, in Internet and Network Economics, Springer, 2007, 306–311.
[19] K. Bhawalkar, S. Gollapudi and K. Munagala, Coevolutionary opinion formation games, STOC’13 Proceedings of the 2013 ACM Symposium on Theory of Computing, 41–50, ACM, New York, 2013.
[20] F. Bonchi, Influence propagation in social networks: A data mining perspective, IEEE Intelligent Informatics Bulletin, 12 (2011), 8–16.
[21] C. Borgs, M. Brautbar, J. Chayes and B. Lucier, Maximizing social influence in nearly optimal time, Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, 946–957, ACM, New York, 2014.

[22] A. Borodin, Y. Filmus and J. Oren, Threshold models for competitive influence in social networks, in Proceedings of the 6th international conference on Internet and network economics, Springer-Verlag, Stanford, CA, USA, 2010, 539–550.

[23] S. Bourigault, C. Lagnier, S. Lamprier, L. Denoyer and P. Gallinari, Learning social network embeddings for predicting information diffusion, WSDM ’14 Proceedings of the 7th ACM International Conference on Web Search and Data Mining, (2014), 393–402.

[24] C. Budak and R. Agrawal, On participation in group chats on twitter, 2013, 165–176.

[25] J. T. Cacioppo, J. H. Fowler and N. A. Christakis, Alone in the crowd: the structure and spread of loneliness in a large social network., Journal of Personality and Social Psychology, 97 (2009), 977.

[26] J. L. Z. Cai, M. Yan and Y. Li, Using crowdsourced data in location-based social networks to explore influence maximization, in Computer Communications, IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on, IEEE, 2016, 1–9.

[27] Z. Cai, Z. He, X. Guan and Y. Li, Collective data-sanitization for preventing sensitive information inference attacks in social networks, IEEE Transactions on Dependable and Secure Computing, (2016), p1.

[28] J. Cannarella and J. A. Spechler, Epidemiological modeling of online social network dynamics, arXiv preprint, arXiv:1401.4208.

[29] T. Carnes, C. Nagarajan, S. M. Wild and A. Van Zuylen, Maximizing influence in a competitive social network: a follower’s perspective, ICEC ’07 Proceedings of the Ninth International Conference on Electronic Commerce, (2007), 351–360.

[30] M. Cha, H. Haddadi, F. Benevenuto and P. K. Gummadi, Measuring user influence in twitter: The million follower fallacy, ICWSM, 10 (2010), 10–17.

[31] M. Cha, A. Mislove and K. P. Gummadi, A measurement-driven analysis of information propagation in the flickr social network, 2009, 721–730.

[32] M. Cha, A. Mislove and K. P. Gummadi, A measurement-driven analysis of information propagation in the flickr social network, in Proceedings of the 18th International Conference on World Wide Web, ACM, Madrid, Spain, 2009, 721–730.

[33] Y. Chang, X. Wang, Q. Mei and Y. Liu, Towards twitter context summarization with user influence models, in Proceedings of the Sixth ACM International Conference on Web Search and Data Mining, ACM, Rome, Italy, 2013, 527–536.

[34] V. Chaoji, S. Ranu, R. Rastogi and R. Bhatt, Recommendations to boost content spread in social networks, in Proceedings of the 21st International Conference on World Wide Web, ACM, Lyon, France, 2012, 529–538.

[35] L. Chen, X. Li and J. Han, Medrank: discovering influential medical treatments from literature by information network analysis, in Proceedings of the Twenty-Fourth Australasian Database Conference, Australian Computer Society, Inc., Adelaide, Australia, 2013, 3–12.

[36] S. Chen, J. Fan, G. Li, J. Feng, K.-I. Tan and J. Tang, Online topic-aware influence maximization, Proceedings of the VLDB Endowment, 8 (2015), 666–677.

[37] W. Chen, A. Collins, R. Cummings, T. Ke, Z. Liu, D. Rincon, X. Sun, Y. Wang, W. Wei and Y. Yuan, Influence maximization in social networks when negative opinions may emerge and propagate, Proceedings of the 2011 SIAM International Conference on Data Mining, 2011, 379–390.

[38] W. Chen, T. Lin and C. Yang, Efficient topic-aware influence maximization using preprocessing, CoRR, abs/1403.0057.

[39] W. Chen, Z. Liu, X. Sun and Y. Wang, A game-theoretic framework to identify overlapping communities in social networks, Data Min. Knowl. Discov., 21 (2010), 224–240.

[40] W. Chen, P. Lu, X. Sun, B. Tang, Y. Wang and Z. A. Zhu, Optimal pricing in social networks with incomplete information, in Internet and Network Economics, Springer, 2011, 49–60.

[41] W. Chen, W. Lu and N. Zhang, Time-critical influence maximization in social networks with time-delayed diffusion process, 2012.

[42] W. Chen, C. Wang and Y. Wang, Scalable influence maximization for prevalent viral marketing in large-scale social networks, in Data Min. Knowl. Discov., 25 (2012), 545–576.

[43] W. Chen, Y. Wang and S. Yang, Efficient influence maximization in social networks, in Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, Paris, France, 2009, 199–208.
[44] W. Chen, Y. Yuan and L. Zhang, Scalable influence maximization in social networks under the linear threshold model, in Proceedings of the 2010 IEEE International Conference on Data Mining, IEEE Computer Society, 2010, 88–97.

[45] Y.-C. Chen, W.-Y. Zhu, W.-C. Peng, W.-C. Lee and S.-Y. Lee, Cim: community-based influence maximization in social networks, ACM Transactions on Intelligent Systems and Technology (TIST), 5 (2014), Article No. 25.

[46] S. Cheng, H. Shen, J. Huang, W. Chen and X. Cheng, Imrank: Influence maximization via finding self-consistent ranking, SIGIR ’14 Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval, (2014), 475–484.

[47] N. A. Christakis and J. H. Fowler, The spread of obesity in a large social network over 32 years, N Engl J Med, 357 (2007), 370–379.

[48] N. A. Christakis and J. H. Fowler, The collective dynamics of smoking in a large social network, New England Journal of Medicine, 358 (2008), 2249–2258.

[49] P. Clifford and A. Sudbury, A model for spatial conflict, Biometrika, 60 (1973), 581–588.

[50] L. Corazzini, F. Pavesi, B. Petrovich and L. Stanca, Influential listeners: An experiment on persuasion bias in social networks, European Economic Review, 56 (2012), 1276–1288.

[51] D. Cosley, D. P. Huttenlocher, J. M. Kleinberg, X. Lan and S. Suri, Sequential influence models in social networks., ICWSM, 10 (2010), 26.

[52] D. M. Cutler and E. L. Glaeser, Social interactions and smoking, Technical report, National Bureau of Economic Research, (2007), 1–28.

[53] A. Das, S. Gollapudi and K. Munagala, Modeling opinion dynamics in social networks, WSDM ’14 Proceedings of the 7th ACM International Conference on Web Search and Data Mining, (2014), 403–412.

[54] A. Das, S. Gollapudi, R. Panigrahy and M. Salek, Debiasing social wisdom, 2013, 500–508.

[55] A. Das, A. Datta, A. D. Procaccia and Y. Zick, Influence in classification via cooperative game theory, arXiv preprint, arXiv:1505.00036.

[56] L. Corazzini, F. Pavesi, B. Petrovich and L. Stanca, Influential listeners: An experiment on persuasion bias in social networks, European Economic Review, 56 (2012), 1276–1288.

[57] E. D. Demaine, M. Hajiaghayi, H. Mahini, D. L. Malec, S. Raghavan, A. Sawant and M. Zadimoghadam, How to influence people with partial incentives, in World Wide Web Conferences, International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 2014, 937–948.

[58] T. N. Dinh, D. T. Nguyen and M. T. Thai, Cheap, easy, and massively effective viral marketing in social networks: truth or fiction?, in Proceedings of the 23rd ACM conference on Hypertext and Social Media, ACM, Milwaukee, Wisconsin, USA, 2012, 165–174.

[59] P. S. Dodds, R. Muhamad and D. J. Watts, An experimental study of search in global social networks, Science, 301 (2003), 827–829.

[60] P. Domingos and M. Richardson, Mining the network value of customers, in Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, San Francisco, California, 2001, 57–66.

[61] Y. Dong, R. A. Johnson and N. V. Chawla, Will this paper increase your h-index?: Scientific impact prediction, Machine Learning and Knowledge Discovery in Databases, (2015), 259–263.

[62] Z. Duan, W. Li and Z. Cai, Distributed auctions for task assignment and scheduling in mobile crowdsensing systems, in Distributed Computing Systems (ICDCS), 2017 IEEE 37th International Conference on, IEEE, 2017, 635–644.

[63] Z. Duan, M. Yan, Z. Cai, X. Wang, M. Han and Y. Li, Truthful incentive mechanisms for social cost minimization in mobile crowdsourcing systems, Sensors, 16 (2016), 481.

[64] I. Eleta, Multilingual use of twitter: Social networks and language choice, in ACM Conference on Computer-Supported Cooperative Work and Social Computing, ACM, New York, NY, USA, 2012, 363–366.

[65] E. Even-Dar and A. Shapira, A note on maximizing the spread of influence in social networks, in Internet and Network Economics, Springer, 2007, 281–286.

[66] Y. Fan and C. R. Shelton, Learning continuous-time social network dynamics, 2009, 161–168.

[67] K. Feng, G. Cong, S. S. Bhawmick and S. Ma, In search of influential event organizers in online social networks, SIGMOD ’14 Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data, 2014, 63–74.

[68] J. H. Fowler, N. A. Christakis, Steptoe and D. Roux, Dynamic spread of happiness in a large social network: Longitudinal analysis of the framingham heart study social network, BMJ: British Medical Journal, 23–27.
[69] L. C. Freeman, A set of measures of centrality based on betweenness, Sociometry, 40 (1977), 35–41.
[70] P. J. Giabbanelli, A. Alimadad, V. Dabbaghian and D. T. Finegood, Modeling the influence of social networks and environment on energy balance and obesity, Journal of Computational Science, 3 (2012), 17–27.
[71] A. Goyal, F. Bonchi and L. V. S. Lakshmanan, Learning influence probabilities in social networks, in Proceedings of the Third ACM International Conference on Web Search and Data Mining, ACM, New York, New York, USA, 2010, 241–250.
[72] A. Goyal, F. Bonchi and L. V. S. Lakshmanan, A data-based approach to social influence maximization, Proc. VLDB Endow., 5 (2011), 73–84.
[73] A. Goyal, F. Bonchi, L. V. Lakshmanan and S. Venkatasubramanian, Approximation analysis of influence spread in social networks, arXiv preprint, arXiv:1008.2005.
[74] A. Goyal, F. Bonchi, L. V. Lakshmanan and S. Venkatasubramanian, On minimizing budget and time in influence propagation over social networks, Social Network Analysis and Mining, 3 (2013), 179–192.
[75] A. Goyal and L. V. S. Lakshmanan, Recmax: Exploiting recommender systems for fun and profit, in Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, Beijing, China, 2012, 1294–1302.
[76] A. Goyal, W. Lu and L. V. S. Lakshmanan, Cell++: Optimizing the greedy algorithm for influence maximization in social networks, in Proceedings of the 20th International Conference Companion on World Wide Web, Proceedings of the 20th international conference companion on World wide web, ACM, Hyderabad, India, 2011, 47–48.
[77] A. Goyal, W. Lu and L. V. S. Lakshmanan, Simpath: An efficient algorithm for influence maximization under the linear threshold model, in Proceedings of the 2011 IEEE 11th International Conference on Data Mining, IEEE Computer Society, 2011, 211–220.
[78] S. Goyal and M. Kearns, Competitive contagion in networks, STOC'12 Proceedings of the 2012 ACM Symposium on Theory of Computing, 2012, 759–774.
[79] M. Grabisch and A. Rusinowska, A model of influence in a social network, Theory and Decision, 69 (2010), 69–96.
[80] M. Granovetter, The strength of weak ties, American Journal of Sociology, 78 (1973), 1.
[81] D. Gruhl, R. Guha, D. Liben-Nowell and A. Tomkins, Information diffusion through blogspace, 2004, 491–501.
[82] A. Guille, H. Hacid, C. E. C. Favre and D. A. Zighed, Information diffusion in online social networks: A survey, ACM SIGMOD Record, 42 (2013), 17–28.
[83] B. Han, P. Hui, V. A. Kumar, M. V. Marathe, J. Shao and A. Srinivasan, Mobile data offloading through opportunistic communications and social participation, Mobile Computing, IEEE Transactions on, 11 (2012), 821–834.
[84] B. Han and A. Srinivasan, Your friends have more friends than you do: identifying influential mobile users through random walks, in Proceedings of the thirteenth ACM international symposium on Mobile Ad Hoc Networking and Computing, ACM, Hilton Head, South Carolina, USA, 2012, 5–14.
[85] M. Han, Z. Duan, C. Ai, F. W. Lybarger, Y. Li and A. G. Bourgeois, Time constraint influence maximization algorithm in the age of big data, International Journal of Computational Science and Engineering, 15 (2017), 165–175.
[86] M. Han, Z. Duan and Y. Li, Privacy issues for transportation cyber physical systems, in Secure and Trustworthy Transportation Cyber-Physical Systems, Springer, Singapore, 2017, 67–86.
[87] M. Han, Q. Han, L. Li, J. Li and Y. Li, Maximizing influence in sensed heterogeneous social network with privacy preservation, International Journal of Sensor Networks, 2017, 1–11.
[88] M. Han, J. Li, Z. Cai and Q. Han, Privacy reserved influence maximization in gps-enabled cyber-physical and online social networks, in Social Computing and Networking (SocialCom), 2016 IEEE International Conferences on, IEEE, 2016, 284–292.
[89] M. Han, J. Li and Z. Zou, Finding k close subgraphs in an uncertain graph, Jisuanji Kexue yu Tansuo, 5 (2011), 791–803.
[90] M. Han, L. Li, X. Peng, Z. Hong and M. Li, Information privacy of cyber transportation system: Opportunities and challenges, RIIT ’17 Proceedings of the 6th Annual Conference on Research in Information Technology, (2017), 23–28.
[91] M. Han, L. Li, Y. Xie, J. Wang, Z. Duan, J. Li and M. Yan, Cognitive approach for location privacy protection, IEEE Access, 6 (2018), 13466–13477.
[92] M. Han, Y. Liang, Z. Duan and Y. Wang, Mining public business knowledge: A case study in sec’s edgar, in Social Computing and Networking (SocialCom), 2016 IEEE International Conferences on, IEEE, 2016, 393–400.

[93] M. Han, J. Wang, M. Yan, C. Ai, Z. Duan and Z. Hong, Near-complete privacy protection: Cognitive optimal strategy in location-based services, Procedia Computer Science, 129 (2018), 298–304.

[94] M. Han, M. Yan, Z. Cai and Y. Li, An exploration of broader influence maximization in timeliness networks with opportunistic selection, Journal of Network and Computer Applications, 63 (2016), 39–49.

[95] M. Han, M. Yan, Z. Cai, Y. Li, X. Cai and J. Yu, Influence maximization by probing partial communities in dynamic online social networks, Transactions on Emerging Telecommunications Technologies, 28 (2017), e3054.

[96] M. Han, M. Yan, J. Li, S. Ji and Y. Li, Generating uncertain networks based on historical network snapshots, in COCOON, 2013, 747–758.

[97] M. Han, M. Yan, J. Li, S. Ji and Y. Li, Neighborhood-based uncertainty generation in social networks, Journal of Combinatorial Optimization, 28 (2014), 561–576.

[98] M. Han, W. Zhang and J.-Z. Li, Raking: An efficient k-maximal frequent pattern mining algorithm on uncertain graph database, Jisuanji Xuebao(Chinese Journal of Computers), 33 (2010), 1387–1395.

[99] R. A. Hanneman and M. Riddle, Introduction to social network methods, 2005.

[100] D. Hatano, T. Fukunaga, T. Maehara and K.-i. Kawarabayashi, Lagrangian decomposition algorithm for allocating marketing channels, 2015.

[101] J. He, J. Hopcroft, H. Liang, S. Suwajanakorn and L. Wang, Detecting the structure of social networks using ($\alpha, \beta$)-communities, in Algorithms and Models for the Web Graph, Springer, 6732 (2011), 26–37.

[102] X. He and D. Kempe, Price of anarchy for the n-player competitive cascade game with submodular activation functions, in Web and Internet Economics, Springer, 2013, 232–248.

[103] X. He and D. Kempe, Stability of influence maximization, KDD ’14 Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, (2014), 1256–1265.

[104] X. He, G. Song, W. Chen and Q. Jiang, Influence blocking maximization in social networks under the competitive linear threshold model, 2012, 463–474.

[105] Z. He, Z. Cai and X. Wang, Modeling propagation dynamics and developing optimized countermeasures for rumor spreading in online social networks, in Distributed Computing Systems (ICDCS), 2015 IEEE 35th International Conference on, IEEE, 2015, 205–214.

[106] Z. He, Z. Cai, J. Yu, X. Wang, Y. Sun and Y. Li, Cost-efficient strategies for restraining rumor spreading in mobile social networks, IEEE Transactions on Vehicular Technology, 66 (2017), 2789–2800.

[107] M. Heidari, M. Asadpour and H. Faiili, Sng: Fast scalable greedy algorithm for influence maximization in social networks, Physica A: Statistical Mechanics and its Applications, 420 (2015), 124–133.

[108] C. Hoede and R. R. Bakker, A theory of decisional power, Journal of Mathematical Sociology, 8 (1982), 309–322.

[109] J. Hopcroft, T. Lou and J. Tang, Who will follow you back?: Reciprocal relationship prediction, CIKM ’11 Proceedings of the 20th ACM International Conference on Information and Knowledge Management, (2011), 1137–1146.

[110] J. Hu, K. Meng, X. Chen, C. Lin and J. Huang, Analysis of influence maximization in large-scale social networks, SIGMETRICS Perform. Eval. Rev., 41 (2014), 78–81.

[111] X. Hu, L. Tang, J. Tang and H. Liu, Exploiting social relations for sentiment analysis in microblogging, in Proceedings of the Sixth ACM International Conference on Web Search and Data Mining, ACM, Rome, Italy, 2013, 537–546.

[112] Z. Hu, J. Yao, B. Cui and E. Xing, Community level diffusion extraction, SIGMOD ’15 Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data, 2015, 1555–1569.

[113] H. Huang, J. Tang, S. Wu, L. Liu and Others, Mining triadic closure patterns in social networks, WWW ’14 Companion Proceedings of the 23rd International Conference on World Wide Web, 2014, 499–504.

[114] J.-P. Huang, C.-Y. Wang and H.-Y. Wei, Strategic information diffusion through online social networks, in Proceedings of the 4th International Symposium on Applied Sciences in
[115] J. Huang, X.-Q. Cheng, H.-W. Shen, T. Zhou and X. Jin, Exploring social influence via posterior effect of word-of-mouth recommendations, in Proceedings of the Fifth ACM International Conference on Web Search and Data Mining, ACM, Seattle, Washington, USA, 2012, 573–582.

[116] J. E. L. Iribarren and E. Moro, Impact of human activity patterns on the dynamics of information diffusion, Physical Review Letters, 103 (2009), 038702.

[117] J. H. Janssen, W. A. IJsselsteijn and J. H. Westerink, How affective technologies can influence intimate interactions and improve social connectedness, International Journal of Human-Computer Studies, 72 (2014), 33–43.

[118] S. Ji, Z. Cai, M. Han and R. Beyah, Whitespace measurement and virtual backbone construction for cognitive radio networks: From the social perspective, in Sensing, Communication, and Networking (SECON), 2015 12th Annual IEEE International Conference on, IEEE, 2015, 435–443.

[119] F. Jiang, S. Jin, Y. Wu and J. Xu, A uniform framework for community detection via influence maximization in social networks, IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014), 2014.

[120] A. P. Joshi, M. Han and Y. Wang, A survey on security and privacy issues of blockchain technology, Mathematical Foundations of Computing, 1 (2018), 121–147.

[121] K. Jung, W. Heo and W. Chen, Irie: A scalable influence maximization algorithm for independent cascade model and its extensions, arXiv preprint, arXiv:1111.4795.

[122] D. Kempe, J. Kleinberg, S. Oren and A. Slivkins, Selection and influence in cultural dynamics, in Proceedings of the fourteenth ACM conference on Electronic commerce, ACM, Philadelphia, Pennsylvania, USA, 2013, 585–586.

[123] D. Kempe, J. Kleinberg and V. Tardos, Influential nodes in a diffusion model for social networks, Automata, Languages and Programming, 2005, 1127–1138.

[124] D. Kempe, J. Kleinberg and E. V. Tardos, Influential nodes in a diffusion model for social networks, Theory Comput., 11 (2015), 105–147.

[125] S. Khanna and B. Lucier, Influence maximization in undirected networks, Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, 1482–1496, ACM, New York, 2014.

[126] Y. A. Kim and J. Srivastava, Impact of social influence in e-commerce decision making, in Proceedings of the Ninth International Conference on Electronic Commerce, ACM, Minneapolis, MN, USA, 2007, 293–302.

[127] M. Kimura, K. Saito, R. Nakano and H. Motoda, Extracting influential nodes on a social network for information diffusion, Data Min. Knowl. Discov., 20 (2010), 70–97.

[128] F. Kooti, W. A. Mason, K. P. Gummadi and M. Cha, Predicting emerging social conventions in online social networks, in Proceedings of the 21st ACM International Conference on Information and Knowledge Management, ACM, Maui, Hawaii, USA, 2012, 445–454, 2390820.

[129] J. Kostka, Y. A. Oswald and R. Wattenhofer, Word of mouth: Rumor dissemination in social networks, in Structural Information and Communication Complexity, Springer, 5058 (2008), 185–196.

[130] R. Kumar, J. Novak, P. Raghavan and A. Tomkins, On the bursty evolution of blogspace, WWW ’03 Proceedings of the 12th international conference on World Wide Web, (2003), 568–576.

[131] R. Kumar, J. Novak and A. Tomkins, Structure and evolution of online social networks, KDD ’06 Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2006, 611–617.

[132] O. Kwon and Y. Wen, An empirical study of the factors affecting social network service use, Computers in Human Behavior, 26 (2010), 254–263.

[133] T. La Fond and J. Neville, Randomization tests for distinguishing social influence and homophily effects, WWW ’10 Proceedings of the 19th International Conference on World Wide Web, (2010), 601–610.

[134] M. Lahiri, A. S. Maiya, R. Sulo, Habiba and T. Y. B. Wolf, The impact of structural changes on predictions of diffusion in networks, in Proceedings of the 2008 IEEE International Conference on Data Mining Workshops, IEEE Computer Society, 2008, 939–948.
[135] X. N. Lam, T. Vu, T. D. Le and A. D. Duong, Addressing cold-start problem in recommendation systems, *ICUIMC ’08 Proceedings of the 2nd International Conference on Ubiquitous Information Management and Communication*, 2008, 208–211.

[136] I. Leftheriotis and M. N. Giannakos, Using social media for work: Losing your time or improving your work?, *Computers in Human Behavior*, **31** (2014), 134–142.

[137] S. Lei, S. Maniu, L. Mo, R. Cheng and P. Senellart, Online influence maximization, *KDD ’15 Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, (2015), 645–654.

[138] J. Leskovec, L. Backstrom and J. Kleinberg, Meme-tracking and the dynamics of the news cycle, *KDD ’09 Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2009, 497–506.

[139] J. Leskovec, M. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen and N. Glance, Cost-effective outbreak detection in networks, in *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, San Jose, California, USA, 2007, 420–429.

[140] J. Leskovec, M. McGlohon, C. Faloutsos, N. Glance and M. Hurst, Information propagation and network evolution on the web, *DA Project, Machine Learning Dept. Carnegie Mellon University*.

[141] J. Leskovec, M. McGlohon, C. Faloutsos, N. S. Glance and M. Hurst, Patterns of cascading behavior in large blog graphs, *Proceedings of the 2007 SIAM International Conference on Data Mining*, **7** (2007), 551–556.

[142] K. Lewis, M. Gonzalez and J. Kaufman, Social selection and peer influence in an online social network, *Proc Natl Acad Sci U S A*, **109** (2012), 68–72.

[143] C.-T. Li, H.-P. Hsieh, S.-D. Lin and M.-K. Shan, Finding influential seed successors in social networks, in *Proceedings of the 21st International Conference Companion on World Wide Web*, ACM, Lyon, France, 2012, 557–558.

[144] D. Li, J. Tang, Y. Ding, X. Shuai, T. Chambers, G. Sun, Z. Luo and J. Zhang, Topic-level opinion influence model (tomin): An investigation using tencent microblogging, *Journal of the Association for Information Science and Technology*, **66** (2015), 2657–2673.

[145] G. Li, S. Chen, J. Feng, K.-A. Tan and W.-s. Li, Efficient location-aware influence maximization, *SIGMOD ’14 Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data*, 2014, 87–98.

[146] H. Li, S. S. Bhowmick, A. Sun and J. Cui, Conformity-aware influence maximization in online social networks, *The VLDB Journal-The International Journal on Very Large Data Bases*, **24** (2015), 117–141.

[147] J. Li, Z. Cai, J. Wang, M. Han and Y. Li, Truthful incentive mechanisms for geographical position conflicting mobile crowdsensing systems, *IEEE Transactions on Computational Social Systems*, **5** (2018), 324–334.

[148] J. Li, X. Guo, L. Guo, S. Ji, M. Han and Z. Cai, Optimal routing with scheduling and channel assignment in multi-power multi-radio wireless sensor networks, *Ad Hoc Networks*, **31** (2015), 45–62.

[149] R.-H. Li, L. Qin, J. X. Yu and R. Mao, Influential community search in large networks, *Proceedings of the VLDB Endowment*, **8** (2015), 509–520.

[150] R. Li, S. Wang, H. Deng, R. Wang and K. C.-C. Chang, Towards social user profiling: Unified and discriminative influence model for inferring home locations, in *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, Beijing, China, 2012, 1023–1031.

[151] Y. Li, W. Chen, Y. Wang and Z.-L. Zhang, Influence diffusion dynamics and influence maximization in social networks with friend and foe relationships, in *Proceedings of the sixth ACM international conference on Web search and data mining*, ACM, Rome, Italy, 2013, 657–666.

[152] Y. Li, D. Zhang and K.-L. Tan, Real-time targeted influence maximization for online advertisements, *Proceedings of the VLDB Endowment*, **8** (2015), 1070–1081.

[153] S.-C. Lin, S.-D. Lin and M.-S. Chen, A learning-based framework to handle multi-round multi-party influence maximization on social networks, *KDD ’15 Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2015, 695–704.

[154] Y. Lin and J. Lui, Algorithmic design for competitive influence maximization problems, *arXiv preprint, arXiv:1410.8664*. 
[155] X. Ling, C. Wu, S. Ji and M. Han, H2dos: An application-layer dos attack towards http/2 protocol, in *Proceedings of SecureComm: Security and Privacy in Communication Networks 2017*, SecureComm ’17, 2017.

[156] B. Liu, G. Cong, Y. Zeng, D. Xu and Y. M. Chee, Influence spreading path and its application to the time constrained social influence maximization problem and beyond, *Knowledge and Data Engineering, IEEE Transactions on*, 26 (2014), 1904–1917.

[157] L. Liu, J. Tang, J. Han, M. Jiang and S. Yang, Mining topic-level influence in heterogeneous networks, in *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, ACM, Toronto, ON, Canada, 2010, 199–208.

[158] L. Liu, J. Tang, J. Han and S. Yang, Learning influence from heterogeneous social networks, *Data Mining and Knowledge Discovery*, 25 (2012), 511–544.

[159] Q. Liu, B. Xiang, E. Chen, H. Xiong, F. Tang and J. X. Yu, Influence maximization over large-scale social networks: A bounded linear approach, *CIKM ’14 Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*, 2014, 171–180.

[160] X. Liu, M. Li, S. Li, S. Peng, X. Liao and X. Lu, Imgpu: Gpu accelerated influence maximization in large-scale social networks.

[161] X. Liu, S. Li, X. Liao, L. Wang and Q. Wu, In-time estimation for influence maximization in large-scale social networks, in *SNS ’12 Proceedings of the Fifth Workshop on Social Network Systems*, 2012, Article No. 3, 1–6.

[162] T. Lou and J. Tang, Mining structural hole spanners through information diffusion in social networks, 2013, 825–836.

[163] T. Lou, J. Tang, J. Hopcroft, Z. Fang and X. Ding, Learning to predict reciprocity and triadic closure in social networks, *ACM Trans. Knowl. Discov. Data*, 7 (2013), 1–25.

[164] J.-L. Lu, L.-Y. Wei and M.-Y. Yeh, Influence maximization in a social network in the presence of multiple influences and acceptances, 2014.

[165] W. Lu, F. Bonchi, A. Goyal and L. V. S. Lakshmanan, The bang for the buck: Fair competitive viral marketing from the host perspective, in *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, Chicago, Illinois, USA, 2013, 928–936.

[166] Y. Lu, J. Ren, J. Qian, M. Han, Y. Huo and T. Jing, Predictive contention window-based broadcast collision mitigation strategy for vanet, in *Social Computing and Networking (SocialCom), 2016 IEEE International Conference on*, IEEE, 2016, 209–215.

[167] Y. Lu, Y. Zhu, M. Han, J. S. He and Y. Zhang, A survey of gpu accelerated svm, in *Proceedings of the 2014 ACM Southeast Regional Conference*, ACM, 2014, Article No. 15.

[168] Z. Lu, L. Fan, W. Wu, B. Thuraisingham and K. Yang, Efficient influence spread estimation for influence maximization under the linear threshold model, *Computational Social Networks*, 1 (2014), 1–19.

[169] B. Lucier, J. Oren and Y. Singer, Influence at scale: Distributed computation of complex contagion in networks, *KDD ’15 Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2015, 735–744.

[170] Z. Z. M Han J Li, K-close: Algorithm for finding the close regions in wireless sensor networks based uncertain graph mining technology, *Journal of Software*, 22 (2011), 131–141.

[171] K. Macropol and A. Singh, Scalable discovery of best clusters on large graphs, *Proceedings of the VLDB Endowment*, 3 (2010), 693–702.

[172] A. S. Maiya and T. Y. Berger-Wolf, Inferring the maximum likelihood hierarchy in social networks, in *Proceedings of the 2009 International Conference on Computational Science and Engineering*, vol. 4, IEEE Computer Society, 2009, 245–250.

[173] M. McPherson, L. Smith-Lovin and J. M. Cook, Birds of a feather: Homophily in social networks, *Annual Review of Sociology*, 27 (2001), 415–444.

[174] I. Mele, F. Bonchi and A. Giovio, The early-adopter graph and its application to web-page recommendation, in *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*, ACM, Maui, Hawaii, USA, 2012, 1682–1686.

[175] S. Mihara, S. Tsugawa and H. Ohsaki, Influence maximization problem for unknown social networks, *ASONAM ’15 Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 2015, 1539–1546.

[176] S. Milgram, The small world problem, *Psychology today*, 2 (1967), 60–67.

[177] A. Mislove, H. S. Koppi, K. P. Gummadi, P. Druschel and B. Bhattacharjee, Growth of the flickr social network, 2008, 25–30.
A. Mislove, M. Marcon, K. P. Gummadi, P. Druschel and B. Bhattacharjee, Measurement and analysis of online social networks, *IMC ’07 Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, 2007, 29–42.

E. Mossel and S. Roch, On the submodularity of influence in social networks, in *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing*, ACM, San Diego, California, USA, 2007, 128–134.

E. Mossel and G. Schoenebeck, Reaching consensus on social networks, 2010, 214–229.

S. A. Myers and J. Leskovec, On the convexity of latent social network inference, *threshold*, 9 (2010), 20.

S. A. Myers and J. Leskovec, The bursty dynamics of the twitter information network, *WWW ’14 Proceedings of the 23rd International Conference on World Wide Web*, 2014, 913–924.

S. A. Myers, C. Zhu and J. Leskovec, Information diffusion and external influence in networks, in *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, Beijing, China, 2012, 33–41.

G. L. Nemhauser, L. A. Wolsey and M. L. Fisher, An analysis of approximations for maximizing submodular set functions. I, *Mathematical Programming*, 14 (1978), 265–294.

M. E. Newman, Spread of epidemic disease on networks, *Physical Review E*, 66 (2002), 016128, 11pp.

M. E. Newman, The structure and function of complex networks, *SIAM Review*, 45 (2003), 167–256.

N. P. Nguyen, T. N. Dinh, X. Ying and M. T. Thai, Adaptive algorithms for detecting community structure in dynamic social networks, *2011 Proceedings IEEE INFOCOM*, 2011.

N. P. Nguyen, G. Yan, M. T. Thai and S. Eidenbenz, Containment of misinformation spread in online social networks, in *Proceedings of the 3rd Annual ACM Web Science Conference*, ACM, Evanston, Illinois, 2012, 213–222.

J. Ok, Y. Jin, J. Choi, J. Shin and Y. Yi, Influence maximization over strategic diffusion in social networks, *2014 48th Annual Conference on Information Sciences and Systems (CISS)*, 2014.

J. P. Onnela and F. Reed-Tsochas, Spontaneous emergence of social influence in online systems, *Proc Natl Acad Sci U S A*, 107 (2010), 18375–18380.

L. Page, S. Brin, R. Motwani and T. Winograd, The pagerank citation ranking: Bringing order to the web.

W. Pan, W. Dong, M. Cebrian, T. Kim, J. H. Fowler and A. S. Pentland, Modeling dynamical influence in human interaction: Using data to make better inferences about influence within social systems, *Signal Processing Magazine, IEEE*, 29 (2012), 77–86.

P. Parchas, F. Gullo, D. Papadias and F. Bonchi, The pursuit of a good possible world: Extracting representative instances of uncertain graphs, *SIGMOD ’14 Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data*, 2014, 967–978.

F. Paulsen, Tönnies, Ferdinando. Gemeinschaft und Gesellschaft. Abhandlungen des communismus und des socialismus als empirischer culturformen. Leipzig, Fues’s Verlag, 1887, Vierteljahresschrift Für Wissenschaftliche Philosophie, 12 (1888), 111–119.

G. Ritzer and Others, *The Blackwell Encyclopedia of Sociology*, vol. 1479, Blackwell Publishing Malden, MA, 2007.

M. G. Rodriguez, D. Baldazzi and B. Sch O Lkopf, Uncovering the temporal dynamics of diffusion networks, arXiv preprint, *arXiv:1105.0697*.

M. G. Rodriguez, J. Leskovec and A. Krause, Inferring networks of diffusion and influence, in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, Washington, DC, USA, 2010, 1019–1028.

M. G. Rodriguez and B. Sch O Lkopf, Influence maximization in continuous time diffusion networks, arXiv preprint, *arXiv:1205.1682*.

D. M. Romero, W. Galuba, S. Asur and B. A. Huberman, Influence and passivity in social media, in *Proceedings of the 20th International Conference Companion on World Wide Web*, ACM, Hyderabad, India, 2011, 113–114.

Y. Rong, X. Wen and H. Cheng, A monte carlo algorithm for cold start recommendation, *WWW ’14 Proceedings of the 23rd International Conference on World Wide Web*, 2014, 327–336.

J. N. Rosenquist, J. H. Fowler and N. A. Christakis, Social network determinants of depression, *Molecular Psychiatry*, 16 (2011), 273–281.
[202] R. A. Rossi, B. Gallagher, J. Neville and K. Henderson, Modeling dynamic behavior in large evolving graphs, in Proceedings of the Sixth ACM International Conference on Web Search and Data Mining, ACM, Rome, Italy, 2013, 667–676.

[203] M. Russell, Mining the Social Web: Analyzing Data from Facebook, Twitter, LinkedIn, and Other Social Media Sites, O’Reilly Media, 2011.

[204] K. Saito, M. Kimura, K. Ohara and H. Motoda, Efficient discovery of influential nodes for sis models in social networks, Knowledge and Information Systems, 30 (2012), 613–635.

[205] K. Saito, M. Kimura, K. Ohara and H. Motoda, Learning asynchronous-time information diffusion models and its application to behavioral data analysis over social networks, Journal of Computer Engineering and Informatics, 1 (2013), 30–57, arXiv:1204.4528.

[206] K. Saito, R. Nakano and M. Kimura, Prediction of information diffusion probabilities for independent cascade model, Knowledge-Based Intelligent Information and Engineering Systems, 5179 (2008), 67–75.

[207] M. Salathé M, E. Kazandjieva, J. W. Lee, P. Levis, M. W. Feldman and J. H. Jones, A high-resolution human contact network for infectious disease transmission, Proceedings of the National Academy of Sciences, 107 (2010), 22020–22025.

[208] D. Sheldon, B. Dilkina, A. N. Elmachtoub, R. Finseth, A. Sabharwal, J. Conrad, C. P. Gomes, D. Shmoys, W. Allen, O. Amundsen and Others, Maximizing the spread of cascades using network design, arXiv preprint, arXiv:1203.3514.

[209] T. Shi, S. Cheng, Z. Cui, Y. Li and J. Li, Retrieving the maximal time-bounded positive influence set from social networks, Personal and Ubiquitous Computing, 20 (2016), 717–730.

[210] H. Shiokawa, Y. Fujiwara and M. Onizuka, Scan++: efficient algorithm for finding clusters, hubs and outliers on large-scale graphs, Proceedings of the VLDB Endowment, 8 (2015), 1178–1189.

[211] X. Shuai, Y. Ding, J. Busemeyer, S. Chen, Y. Sun and J. Tang, Modeling indirect influence on twitter, International Journal on Semantic Web and Information Systems (IJSWIS), 8 (2012), 20–36.

[212] Y. Singer, How to win friends and influence people, truthfully: Influence maximization mechanisms for social networks, in Proceedings of the Fifth ACM International Conference on Web Search and Data Mining, ACM, Seattle, Washington, USA, 2012, 733–742.

[213] R. Sipos, A. Ghosh and T. Joachims, Was this review helpful to you?: It depends! context and voting patterns in online content, WWW ’14 Proceedings of the 23rd International Conference on World Wide Web, 2014, 337–348.

[214] D. Song and D. A. Meyer, A model of consistent node types in signed directed social networks, 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014), 2014.

[215] G. Song, X. Zhou, Y. Wang and K. Xie, Influence maximization on large-scale mobile social network: A divide-and-conquer method, Parallel and Distributed Systems, IEEE Transactions on, 26 (2015), 1379–1392.

[216] J. Stehlé E, N. Voirin, A. Barrat, C. Cattuto, L. Isella, J.-F. C. C. O. Pinton, M. Quaggiotto, W. Van den Broeck, C. R E Gis, B. Lina and Others, High-resolution measurements of face-to-face contact patterns in a primary school, PloS one, 6 (2011), 23176.

[217] J. Sun and J. Tang, A survey of models and algorithms for social influence analysis, in Social Network Data Analytics, Springer, 2011, 177–214.

[218] T. Sun, W. Chen, Z. Liu, Y. Wang, X. Sun, M. Zhang and C.-Y. Lin, Participation maximization based on social influence in online discussion forums, 2011.

[219] F. Tang, Q. Liu, H. Zhu, E. Chen and F. Zhu, Diversified social influence maximization, 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014), 2014.

[220] J. Tang, J. Sun, C. Wang and Z. Yang, Social influence analysis in large-scale networks, in Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, Paris, France, 2009, 807–816, 1557108.

[221] J. Tang, B. Wang, Y. Yang, P. Hu, Y. Zhao, X. Yan, B. Gao, M. Huang, P. Xu, W. Li and Others, Patentminer: topic-driven patent analysis and mining, KDD ’12 Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2012, 1366–1374.

[222] J. Tang, S. Wu and J. Sun, Confluence: Conformity influence in large social networks, in Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, Chicago, Illinois, USA, 2013, 347–355.
[223] J. Tang, C. Zhang, K. Cai, L. Zhang and Z. Su, Sampling representative users from large social networks, 2015.
[224] J. Tang, Y. Zhang, J. Sun, J. Rao, W. Yu, Y. Chen and A. C. M. Fong, Quantitative study of individual emotional states in social networks, Affective Computing, IEEE Transactions on, 3 (2012), 132–144.
[225] X. Tang and C. C. Yang, Ranking user influence in healthcare social media, ACM Trans. Intell. Syst. Technol., 3 (2012), Article No. 73.
[226] Y. Tang, X. Xiao and Y. Shi, Influence maximization: Near-optimal time complexity meets practical efficiency, SIGMOD ’14 Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data, 2014, 75–86.
[227] J. Teevan, D. Ramage and M. R. Morris, Twittersearch: A comparison of microblog search and web search, WSDM ’11 Proceedings of the Fourth ACM International Conference on Web Search and Data Mining, 2011, 35–44.
[228] G. Tong, W. Wu, S. Tang and D.-Z. Du, Adaptive influence maximization in dynamic social networks, IEEE/ACM Transactions on Networking, 25 (2017), 112–125, arXiv:1506.06294.
[229] W. Tang, R. Goebel and G. Lin, Smoothed heights of tries and patricia tries, Theoretical Computer Science, 609 (2016), 620–626.
[230] H. Trottier and P. Philippe, Deterministic modeling of infectious diseases: theory and methods, The Internet Journal of Infectious Diseases, 1 (2001), 3.
[231] J. Tsai, T. H. Nguyen and M. Tambe, Security games for controlling contagion, 2012.
[232] W. Verbeke, D. Martens and B. Baesens, Social network analysis for customer churn prediction, Applied Soft Computing, 14 (2014), 431–446.
[233] J. Videras, A. L. Owen, E. Conover and S. Wu, The influence of social relationships on pro-environmental behaviors, Journal of Environmental Economics and Management, 63 (2012), 35–50.
[234] B. Viswanath, A. Mislove, M. Cha and K. P. Gummadi, On the evolution of user interaction in facebook, WSN’09 Proceedings of the 2nd ACM Workshop on Online Social Networks, 2009, 37–42.
[235] R. W O Lfer and H. Scheithauer, Social influence and bullying behavior: Intervention-based network dynamics of the fairplayer. manual bullying prevention program, Aggressive behavior.
[236] C. Wang, W. Chen and Y. Wang, Scalable influence maximization for independent cascade model in large-scale social networks, Data Mining and Knowledge Discovery, 25 (2012), 545–576.
[237] F. Wang, E. Camacho and K. Xu, Positive influence dominating set in online social networks, in Proceedings of the 3rd International Conference on Combinatorial Optimization and Applications, Springer-Verlag, Huangshan, China, 5573 (2009), 313–321.
[238] G. Wang, Q. Hu and P. S. Yu, Influence and similarity on heterogeneous networks, in Proceedings of the 21st ACM International Conference on Information and Knowledge Management, ACM, Maui, Hawaii, USA, 2012, 1462–1466.
[239] Y. Wang, G. Cong, G. Song and K. Xie, Community-based greedy algorithm for mining top-k influential nodes in mobile social networks, in Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, Washington, DC, USA, 2010, 1039–1048.
[240] D. J. Watts and S. H. Strogatz, Collective dynamics of “small-world” networks, The Structure and Dynamics of Networks, (2011), 301–303.
[241] J. Weng, E. P. Lim, J. Jiang and Q. He, Twittrank: finding topic-sensitive influential twitters, WSDM ’10 Proceedings of the Third ACM International Conference on Web Search and Data Mining, 2010, 261–270.
[242] C. Wilson, A. Sala, K. P. N. Puttaswamy and B. Y. Zhao, Beyond social graphs: User interactions in online social networks and their implications, ACM Trans. Web, 6 (2012), 1–31.
[243] M. Workman, New media and the changing face of information technology use: The importance of task pursuit, social influence, and experience, Computers in Human Behavior, 31 (2014), 111–117.
[244] S. Wu, J. Sun and J. Tang, Patent partner recommendation in enterprise social networks, in Proceedings of the Sixth ACM International Conference on Web Search and Data Mining, ACM, Rome, Italy, 2013, 43–52.
[245] X. Xu, N. Yuruk, Z. Feng and T. A. J. Schweiger, Scan: a structural clustering algorithm for networks, KDD ’07 Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2007, 824–833.

[246] M. Yan, M. Han, C. Ai, Z. Cai and Y. Li, Data aggregation scheduling in probabilistic wireless networks with cognitive radio capability, in IEEE GLOBECOM 2016, 2016.

[247] M. Yan, S. Ji, M. Han, Y. Li and Z. Cai, Data aggregation scheduling in wireless networks with cognitive radio capability, in Sensing, Communication, and Networking (SECON), 2014 Eleventh Annual IEEE International Conference on, IEEE, 2014, 513–521.

[248] D.-N. Yang, H.-J. Hung, W.-C. Lee and W. Chen, Maximizing acceptance probability for active friending in on-line social networks, arXiv preprint, arXiv:1302.7025.

[249] Y. Yang, J. Jia, S. Zhang, B. Wu, Q. Chen, J. Li, C. Xing and J. Tang, How do your friends on social media disclose your emotions?, 2014.

[250] Y. Yang, J. Tang, C. Leung, Y. Sun, Q. Chen, J. Li and Q. Yang, Rain: Social role-aware information diffusion, 2015.

[251] Z. Yang, J. Tang, B. Xu and C. Xing, Active learning for networked data based on non-progressive diffusion model, WSDM ’14 Proceedings of the 7th ACM International Conference on Web Search and Data Mining, 2014, 363–372.

[252] H. Yoganarasimhan, Impact of social network structure on content propagation: A study using youtube data, Quantitative Marketing and Economics, 10 (2012), 111–150.

[253] H. Yu, S.-K. Kim and J. Kim, Scalable and parallelizable processing of influence maximization for large-scale social networks?, in Proceedings of the 2013 IEEE International Conference on Data Engineering (ICDE 2013), IEEE Computer Society, 2013, 266–277.

[254] X. Yu, X. Ren, Y. Sun, B. Sturt, U. Khandelwal, Q. Gu, B. Norick and J. Han, Recommendation in heterogeneous information networks with implicit user feedback, RecSys ’13 Proceedings of the 7th ACM Conference on Recommender Systems, 2013, 347–350.

[255] Y. Yu, T. Y. Berger-Wolf, J. Saia and Others, Finding spread blockers in dynamic networks, in Advances in Social Network Mining and Analysis, Springer, 2010, 55–76.

[256] H. Zhang, A. D. Procaccia and Y. Vorobeychik, Dynamic influence maximization under increasing returns to scale, 2015.

[257] H. Zhang, T. N. Dinh and M. T. Thai, Maximizing the spread of positive influence in online social networks, 2013 IEEE 33rd International Conference on Distributed Computing Systems, 2013.

[258] H. Zhang, S. Mishra and M. T. Thai, Recent advances in information diffusion and influence maximization of complex social networks, Opportunistic Mobile Social Networks, 37.

[259] J. Zhang, B. Liu, J. Tang, T. Chen and J. Li, Social influence locality for modeling retweeting behaviors, in Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence, AAAI Press, Beijing, China, 2013, 2761–2767.

[260] J. Zhang, J. Tang, C. Ma, H. Tong, Y. Jing and J. Li, Panther: Fast top-k similarity search in large networks, KDD ’15 Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2015, 1445–1454, arXiv:1504.02577.

[261] J. Zhang, J. Tang, H. Zhuang, C. W.-K. Leung and J. Li, Role-aware conformity influence modeling and analysis in social networks, 2014.

[262] M. Zhang, J. Tang, X. Zhang and X. Xue, Addressing cold start in recommender systems: A semi-supervised co-training algorithm, SIGIR ’14 Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval, 2014, 73–82.

[263] P. Zhang, W. Chen, X. Sun, Y. Wang and J. Zhang, Minimizing seed set selection with probabilistic coverage guarantee in a social network, KDD ’14 Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2014, 1306–1315.

[264] J. Zhao, J. Wu, X. Feng, H. Xiong and K. Xu, Information propagation in online social networks: A tie-strength perspective, Knowledge and Information Systems, 32 (2012), 589–608.

[265] C. Zhou, P. Zhang, W. Zang and L. Guo, Maximizing the cumulative influence through a social network when repeat activation exists, Procedia Computer Science, 29 (2014), 422–431.

[266] Y. Zhou and L. Liu, Social influence based clustering of heterogeneous information networks, in Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, Chicago, Illinois, USA, 2013, 338–346.
[267] H. Zhu, B. Huberman and Y. Luon, *To switch or not to switch: Understanding social influence in online choices*, in *CHI ’12*, CHI ’12, ACM, New York, NY, USA, 2012, 2257–2266.

[268] H. Zhuang, Y. Sun, J. Tang, J. Zhang and X. Sun, *Influence maximization in dynamic social networks*, 2013 IEEE 13th International Conference on Data Mining, 2013, 1313–1318.

Received December 2017; revised February 2018.

E-mail address: menghan@kennesaw.edu
E-mail address: yili@gsu.edu