

Social Influence Analysis in Social Networking Big Data: Opportunities and Challenges

Sancheng Peng, Guojun Wang, and Dongqing Xie

ABSTRACT

Social influence analysis has become one of the most important technologies in modern information and service industries. It will definitely become an essential mechanism to perform complex analysis in social networking big data. It is attracting an increasing amount of research ranging from popular topics extraction to social influence analysis, including analysis and processing of big data, social influence evaluation, influential users identification, and information diffusion modeling. We provide a comprehensive investigation of social influence analysis, and discuss the characteristics of social influence and the architecture of social influence analysis based on social networking big data. The relationship between big data and social influence analysis is also discussed. In addition, research challenges relevant to real-world issues based on social networking big data in social influence analysis are discussed, focusing on research issues such as scalability, data collection, dynamic evolution, causal relationships, network heterogeneity, evaluation metrics, and effective mechanisms. Our goal is to provide a broad research guideline of existing and ongoing efforts via social influence analysis in large-scale social networks, and to help researchers better understand the existing work, and design new algorithms and methods for social influence analysis.

INTRODUCTION

Social networking is used widely as an important communication media with exponential growth. Specifically, Web 2.0 technology has brought revolutionary changes to our daily lives in online social networks. In the last decade, lots of online social networks, such as Facebook, Twitter, and LinkedIn, have emerged and tightly connected web users all over the world. People can directly engage in these networks, then build their own friendship networks, and share their opinions, insights, information, experiences, and perspectives with each other. In addition, they can discover and spread information by using various formats such as words, pictures, audio, and video.

Social networking big data is a collection of very huge data sets with a great diversity of types from social networks (e.g., Facebook, Twitter, and LinkedIn). The 5V characteristics of social networking big data are volume, velocity, variety, value, and veracity, which make it difficult to handle such big data using traditional tools and techniques. Nowadays, the explosion of data in terms of high volume, high velocity, and high vari-

ety, fueled by stunning and exciting advances in information technologies and web techniques, has become the focus of widespread attention. Big data applications [1] lie in many scientific disciplines, such as physics, astronomy, atmospheric science, medicine, genomics, biology, biogeochemistry, and other complex and interdisciplinary areas of scientific research.

Social influence analysis [2] is pervasive throughout society. The high-level goal of social influence analysis is to answer questions related to social influence, such as "Who can be influenced?," "Who can influence whom?," "Who are vulnerable to influence?," "Why is a user attracted to a particular group?," and "Who are the most influential users in a specific social network?" The main idea of social influence analysis is how to quantify the influence of each user and how to identify the most influential users in social networks. It has great potential to help us understand the ways in which information, experiences, ideas, and innovations propagate across social networks. Analyzing and mining social networks can provide new insights into how people interact with and influence each other, and why their ideas and opinions on different subjects can spread in social networks.

Analysis and modeling on social influence is becoming an important part of research on social networks. By analyzing the influencing mode among users and the spreading mode of influence, based on social networking big data, the following advantages can be obtained: understanding of social behaviors of people from the angle of sociology, a theoretical basis for making public decisions and public opinion guidance, and promoting communication and dissemination of political, economic, and cultural activities, as well as in other fields. Thus, social influence analysis in social networks has important social significance and application value using social networking big data.

The goal of this article is to provide a comprehensive investigation of the status of social networking big data for social influence analysis, and the characteristics and architecture of social influence analysis in social networking big data. The relationship between social influence analysis and big data is discussed, and the opportunities for social influence analysis are also discussed. In addition, research challenges are provided, focusing on research issues such as scalability, data collection, dynamic evolution, causal relationship, network heterogeneity, evaluation metrics, and effective mechanisms.

The remainder of this article is organized as

Sancheng Peng is with
Guangdong University of
Foreign Studies.

Guojun Wang is with
Guangzhou University.

Dongqing Xie is with
Guangzhou University.

Digital Object Identifier:
10.1109/MNET.2016.1500104NM

Due to the rapid development of the Internet, Internet of Things, and cloud computing, data volume is increasing faster than during any previous period, data size is getting larger, and data types are becoming more diverse and complex, which means that social networking big data is penetrating into all aspects of human society, and the era of social networking big data is emerging unannounced.

follows. We describe social influence analysis based on social networking big data, and provide opportunities for social influence analysis. We discuss challenges for social influence analysis and conclude this article.

SOCIAL INFLUENCE ANALYSIS IN SOCIAL NETWORKING BIG DATA

THE RELATIONSHIP BETWEEN SOCIAL INFLUENCE ANALYSIS AND BIG DATA

Due to the rapid development of the Internet, the Internet of Things, and cloud computing, data volume is increasing faster than during any previous period, data size is getting larger, and data types are becoming more diverse and complex, which means that social networking big data is penetrating all aspects of human society, and the era of social networking big data is emerging unannounced.

However, social networking big data still remains both an opportunity and a challenge for social influence analysis. By combining information technology and social networking theory, the huge amounts of data can be completely used to address the challenges in various application fields, such as viral marketing, stock market prediction, domain expert finding, correlation with the incidence of influenza, and influential user identification. To meet these challenges, we discuss the opportunities and challenges of social influence analysis in social networking big data in this article.

Social influence [3, 4] refers to when individuals change their behaviors under the influence of others. The strength of social influence depends on the relation among individuals, network distances, time effect, characteristics of networks and individuals, and so on. Viral marketing, online advertising, recommendations, and other applications can benefit from social influence by qualitatively and quantitatively measuring the influence of individuals on others or the influence of one group on others.

Social influence analysis is a highly utilizable technology, which is attracting a large number of researchers to develop it and solve real-world issues based on social networking big data. The increasing development of large-scale social networks is promoting the emergence of new data analytic tools and algorithms for big data techniques. Thus, these analytic tools and algorithms facilitate scalable, accessible, and sustainable data infrastructure to increase understanding of human and social processes and interactions.

Online social networking services, such as Facebook, Twitter, and LinkedIn, have become increasingly popular over the past few years. In general, such services include massive linked data and content data. The linked data is mainly in the form of graphic structures, describing the commu-

nications between any two entities. The content data contains text, images, audio, and video. The rich content in such networks brings about both unprecedented challenges and opportunities for social influence analysis. On one hand, the development of social networks results in the emergence of social networking big data. On the other hand, the emergence of social networking big data also accelerates the development of social influence analysis, cloud computing, and other technologies of big data.

The relationship between social influence analysis and big data is shown in Fig. 1.

PROPERTIES OF SOCIAL INFLUENCE

Social influence [5] is defined as the power exerted by an individual i on an individual j , having the effect of change in the opinion of individual j . The properties of social influence are described as follows.

Dynamic: Influence referring to the impact is not absolute. It changes with change of context and time dynamically.

Propagative: Influence is propagative. Due to its propagative nature, information can be passed from one member to another in a social network, creating influence chains. Spreading influence in a social network is the basis of “word-of-mouth” propagation of information for human beings.

Composable: Propagation of influence along social chains allows a member to form some influence on another member not directly connected to him/her. However, when several chains influence a member indirectly, the influencer needs to compose the influence information.

Measurable: The level of influence can be measured by a continuous real number, referred to as the influence value. Influence can also be represented with uncertainty (e.g., strongest, stronger, strong, medium, weak, weaker, weakest).

Subjective: Influence is subjective. The nature of influence leads to personalization of influence computation, where the biases and preferences of the influencer may have a direct impact on the computed influence value.

Asymmetric: Influence is typically asymmetric. For example, if Bob influences Alice, it does not imply that Alice also influences Bob. A member may influence another member more than he or she is influenced back.

Event-sensitive: Influence may take a long time to build, but a single high-impact event may destroy it completely in a short time.

ARCHITECTURE OF SOCIAL INFLUENCE ANALYSIS

The architecture of social influence analysis in social networking big data is shown in Fig. 2.

Data collection from social networks: It is a very important basis of social influence analysis. With the availability of social networking big data generated by online shopping, advertising, instant messaging, and mobile communication, the collection of raw data from online sources (e.g., Twitter, Facebook) and offline sources (e.g., call data) is much easier.

Data preprocessing: In order to improve the performance and convenience of processing, we need to remove irrelevant information from social influence analysis. In order to protect the privacy of users, we also need to filter out sensitive

information related to privacy protection from the collected data sets.

Selection of evaluation metrics: It is very important to extract a set of evaluation metrics to accurately characterize the characteristics of each user, as these evaluation metrics are helpful to quantify the social influence of each user and to easily find the most influential top- k nodes. In the existing work, common evaluation metrics include the centrality of each node, interaction frequency, and so on.

Modeling and computing social influence: According to the extracted evaluation metrics, an evaluation model and computing equations are provided to a specific social network. Thus, the social influence of each user can be computed by integrating the computing equations into the collected real-world data sets.

Selection of the most influential top- k nodes: The influence maximization algorithm [6] is designed to find the most influential top- k nodes. Most of the existing algorithms are improved algorithms based on the greedy algorithm.

Performance analysis: Simulation is done to validate performance (e.g., influential range, computational complexity) of the proposed algorithms based on a specific propagation model. For an excellent algorithm, it has a large influential range and low computational complexity.

OPPORTUNITIES FOR SOCIAL INFLUENCE ANALYSIS

With the rapid increase of huge data sets generated by various social networks, including online shopping, advertising, mobile communications, instant messaging, Wechat, micro blogs, and so on, a golden opportunity for social influence analysis is coming. Researchers can work on many types of research, such as quantifying the influence of each user, identifying the most influential users, analyzing the behaviors of each user, and characterizing the spreading dynamics of information in the networks. There are many opportunities described as follows.

With the rapid increase of huge data sets generated by various social networks, including online shopping, advertising, mobile communications, instant messaging, Wechat, micro blogs, and so on, a golden opportunity for social influence analysis is coming.

PROMOTING THE DEVELOPMENT OF SOCIAL NETWORKS

With the development of Internet technologies, web users are growing rapidly. For example, Facebook has 1.36 billion users, with five new profiles being created every second and 4.5 billion interests every day. Twitter has a billion users and generates 6000 tweets per second. Relative newcomer Google+ has 540 million active users and gains 800,000 each month. Even the professional network LinkedIn has 300 million users with two more profiles added every second. Thus, some network companies have to face the challenges of handling such social networking big data. Taking Facebook as an example, it has introduced the Hadoop 2.0 platform to manage its server data.

The Hadoop 2.0 platform makes Facebook one of the most relevant and diffused social media platforms. Facebook allows users to connect to their friends and families using a webpage called a profile. The majority of data in Facebook is generated by web users, from their comments, interests, audios, videos, images, and so on. The profile acts as a medium from which to interact with other profiles of members. New content is generated by a user, and members may add their comments to this content or forward this content to other profiles of members. For social networks, this phenomenon is called the “word-of-mouth” effect, which leads to a virtual circle of new content by users.

The advantages of social networks have gathered people from around the world to be a part of them. As part of one’s daily routine, there are many advantages (e.g., convenient, simple, quick, practical, efficient, economical) that influence people to use online social networks.

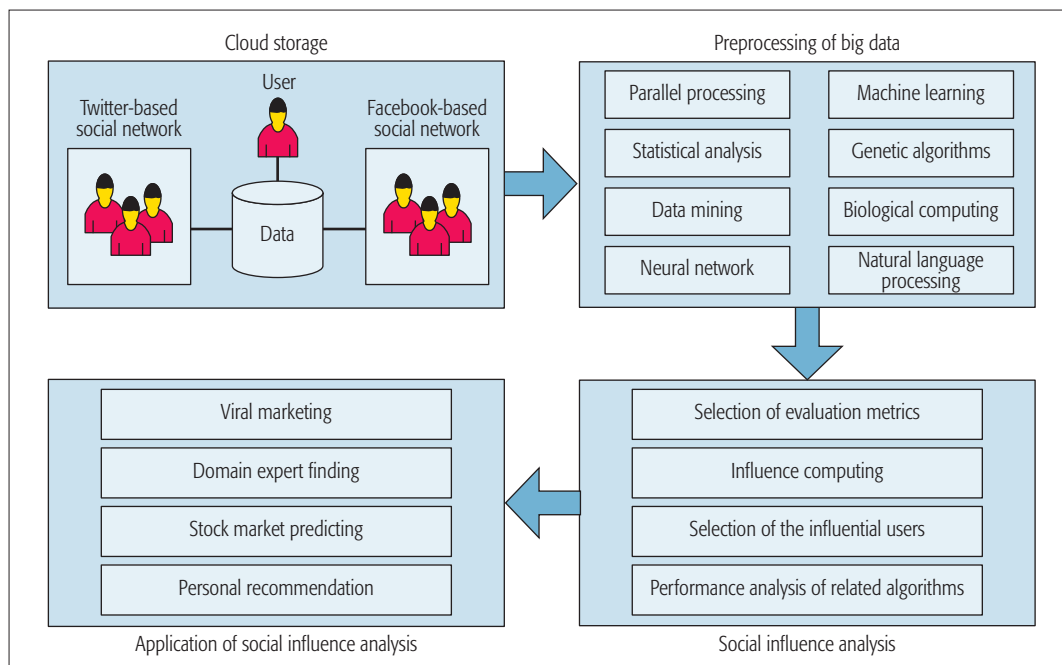


FIGURE 1. The relationship between social influence analysis and big data.

With the enormous volume and variety of information from social media, social networking big data can provide an excellent opportunity for social influence analysis to extract knowledge for particular predictions with specific outcomes.

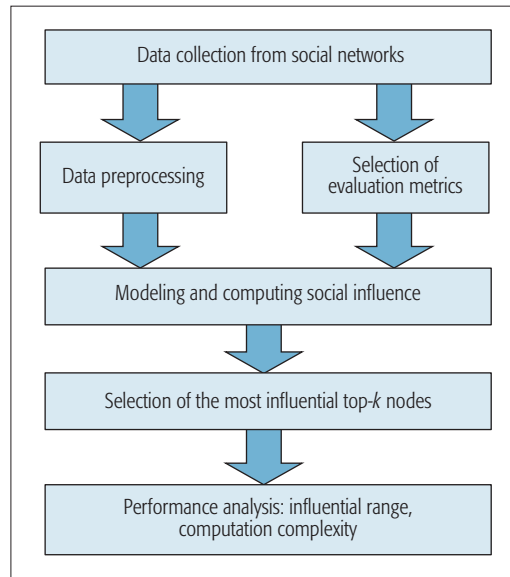


FIGURE 2. Architecture of social influence analysis.

LAYING A SOUND THEORETICAL FOUNDATION FOR RESEARCH ON SOCIAL INFLUENCE ANALYSIS

Increasing advancement in big data technology offers opportunities for improvements in critical decision-making development areas, such as healthcare, employment, economic productivity, crime, security, natural disasters, and resource management [7].

With wider and deeper study on the modeling and analysis of social networking big data, various technologies and methods are being exploited in this domain, such as machine learning, data mining, parallel processing, cloud computing, natural language processing, genetic algorithms, biological computing, statistical analysis, and neural networks. These technologies and methods are paving a broad way to study social influence analysis. For example, [8] introduced a parallel processing algorithm to solve influence maximization.

FACILITATING COLLECTION AND PROCESSING OF SOCIAL NETWORKING BIG DATA

Cloud computing is closely related to social networking big data. It has become a viable, mainstream solution for big data processing, storage, and distribution. Social networking big data is an object of computation-intensive operation and stresses the storage capacity of a cloud system. The goal of cloud computing [3] is to use huge computing and storage resources under concentrated management so as to provide big data applications with fine-grained computing capacity. The development of cloud computing provides a sound solution for the collection and processing of social networking big data.

In addition, the emergence of social networking big data also accelerates the growth of cloud computing. The distributed storage technology

based on cloud computing can effectively manage social networking big data; techniques like parallel computing and data mining by virtue of cloud computing improve the efficiency of analysis on social networking big data.

IMPROVING THE APPLICATION PROSPECTS OF SOCIAL INFLUENCE ANALYSIS

With the enormous volume and variety of information from social media, social networking big data can provide an excellent opportunity for social influence analysis to extract knowledge for particular predictions with specific outcomes. This kind of social media data can be used to build models to aggregate opinions from the collective community and to gain some useful insights into their behaviors that can be used to predict future trends.

In addition, it can be used to gather information on the comments people make with respect to a particular product. Analysis of such comments shows it to be valuable for the design of marketing and advertising campaigns; analysis of the dissemination characteristics of information in social networks can provide research ideas for containing the spreading of negative information in the network; and detection on the sharp growth or drop of the amount of topics can predict the occurrence of abnormal events.

CHALLENGES FOR SOCIAL INFLUENCE ANALYSIS

Social networking big data brings large amounts of attractive opportunities to social influence analysis, and we present a positive outlook for social influence analysis and provide an overview of the technical areas that we consider most relevant to the future various applications. However, opportunities are always followed by challenges. Due to the short history of social influence analysis in social networking big data, related research on this topic is still in its infancy, and there are many challenges highlighted in each area and listed as follows.

DETERMINING A SET OF EFFECTIVE EVALUATION METRICS

Due to the heterogeneity of big data systems, the data sets used in existing work are collected from a variety of venues by various methods. The process of social influence analysis should be diverse and multi-faceted. Thus, it is very necessary to study the evaluation metrics about social influence. A set of effective evaluation metrics would be helpful to effectively quantify social influence, to provide directional guidance for the design of a new evaluation model, and also to describe more accurately the complex phenomena in social networks. However, it is a challenging problem to determine a set of effective evaluation metrics in a real social network, as the scale of social networks is becoming very large, and the topology of social networks is more complex.

CONSIDERING THE CHARACTERISTICS OF DYNAMIC EVOLUTION IN LARGE-SCALE SOCIAL NETWORKS

In the majority of early existing work, it is assumed that a social network is a static network in the study of social influence analysis. However, one important aspect has been neglected, that is, the change of a social network over time. In fact, a given social network is constantly evolving,

and from a practical perspective, being able to react to these changes quickly is very crucial. In addition, the social relationship between any two users is time-varying, and it is sometimes hidden even. This uncertainty of social graph inspires us to study a different kind of social influence analysis. However, research on this domain is still undeveloped, and much work needs to be done in designing evolution models and understanding how such evolution impacts particular goals.

CHARACTERIZING CAUSAL RELATIONSHIPS IN LARGE-SCALE SOCIAL NETWORKS

Social networks have become increasingly attractive to both users and the companies providing these networks. It is essential to ensure efficient management of these networks, which requires knowledge of causal relationships in the networks. Social networks are widely used and intrinsically interesting to analyze due to the following two main reasons. First, these networks allow us to collect massive raw data about social interactions of users. These data sets can be analyzed to identify causal relationships in a way that reveals the user behavior in a social network. Second, causal analysis of this data set can provide a better understanding of the impact of network design, and how users collaborate and interact with one another. Such understanding is important for management of these networks and can provide useful insights in making better design decisions for future versions of these networks.

There are many causal problems [9] to be resolved in social networks (e.g., Blogosphere, Twitter), including “what causes a user to publish a micro-blog?”; “what causes another user to forward a micro-blog?”; “what causes a user to purchase a particular product?”; “what causes information to be propagated among nodes?”; and “what causes a user to click on a particular advertisement?”

Fortunately, some researchers introduced transfer entropy [10] to characterize the causal relationships in social networks, which is a good beginning. However, lots of problems need to be solved in large-scale social networks.

DISTINGUISHING POSITIVE INFLUENCE, NEGATIVE INFLUENCE, AND CONTROVERSY INFLUENCE

There are three types of influence in a social network, including positive influence, negative influence, and controversy influence. Positive influence and negative influence show that an individual has higher influence on others. The principal difference is that the individual’s influence is considered from different perspectives:

- Positive influence represents that an individual has a low negative influence, and its links with other individuals express agreement, trust, support, or approval. The opinions carry influence in a positive manner; that is, a user would be more likely to trust and adopt his/her friends’ opinions.
- Negative influence indicates that an individual has a low positive influence and has connections with other individuals such as disagreement, distrust, hostility, or disapproval. The opinions carry influence in a negative manner; that is, a user would be more likely to distrust and reject his/her friends’ opinions.

It is difficult to guarantee efficiency and scalability on social influence analysis in social networking big data. At the micro level, evaluating influence is a challenge for scalability: influence evaluation for each node occupies most of the processing time. At the macro level, the problem of finding the optimal seed nodes for influence maximization is NP-hard, and also hard to solve in polynomial time.

- Controversy influence indicates that the opinions of an individual are likely to be challenged or accepted by other individuals. The rapport between positive and negative influence is relatively balanced.

However, how to distinguish these three types of influence in a real social network is really a challenge.

GUARANTEEING THE EFFICIENCY AND SCALABILITY OF SOCIAL INFLUENCE ANALYSIS

It is difficult to guarantee efficiency and scalability of social influence analysis in social networking big data. At the micro level, evaluating influence is a challenge for scalability: influence evaluation for each node occupies most of the processing time. At the macro level, the problem of finding the optimal seed nodes for influence maximization is NP-hard [11], and also hard to solve in polynomial time. Although Kempe proposed the greedy algorithm to guarantee constant approximation, the diffusion speed of maximization is still intractable even if the difficulty of evaluation in the diffusion time can be ignored.

Moreover, with the continuous expansion of the scale of online social networks, most existing methods encounter the runtime efficiency problem, and it is difficult to implement them on a large scale. For example, there are hundreds of or thousands of online researchers, and millions of research material data in ScholarMate [12], so finding experts in similar research fields and ranking them in terms of some criterion are time-consuming.

However, the gain in scalability is obtained at the cost of unguaranteed efficiency. In a word, existing algorithms and methods of social influence analysis suffer from the scalability-efficiency dilemma.

EVALUATING THE INFLUENCE OF HETEROGENEOUS SOCIAL NETWORK

Analyzing social influence in heterogeneous social networks [13] with multiple types of entities, links, static or dynamic attributes, and interconnected activities demands new analysis models or methods and is also a new challenge.

- The increasing scale of heterogeneous social networks indicates features of social complexity, and also involves substantial nontrivial computational cost.
- Each type of entity is usually associated with one primary social network but participates in many other social networks, each with domain-specific semantics. How can we make full use of the information from various social networks to provide more informative views on how people influence one another in a specific social network?
- The information flow between two social networks may be bidirectional, so we should be careful in differentiating them when we integrate the results from different social networks.

In real-world online social network services, such as Twitter or Facebook, each user has a distinct spreading rate on propagating information according to a user's reputation and/or role, such as opinion formers, leaders, or followers. It is a challenging direction to design hybrid approaches that combine the advantages of different algorithms to further improve the efficiency and effectiveness of social influence analysis.

- As multiple social networks may be from arbitrary domains, it is challenging to efficiently integrate multiple types of influences from multiple social networks into a unified mechanism for social influence analysis simultaneously.

CONSIDERING THE COLLECTION AND PROCESSING OF SOCIAL NETWORKING BIG DATA

In the early days, since there were no "automatic" methods to collect data in most social scientific research, data collection was done by performing interviews and often by small-scale group studies with volunteers [14], so collection of relevant data was one of the biggest obstacles to research on social influence analysis. Nowadays, with the rapid development of Web 2.0 technology, the collection of raw data from online sources (e.g., Twitter, Facebook) and offline sources (e.g., call data) is much easier to obtain. Considering the growing size of social networks, the urgent task is to develop efficient algorithms to handle these social networking big data.

As mentioned in the previous section, social networking big data is helpful to perform influence analysis. However, data quality is always an important issue, and there are several new challenges, such as the computational complexity in analyzing a social network with millions or billions of nodes, and the integration of multiple data sources with implicit connections. With the availability of social networking big data generated by online shopping, advertising, and instant messaging, how to identify the most influential users for a given target user in multiple data sources is also a challenge. In addition, due to the sensitivity of information on social relationships, additional privacy issues may arise [14]. As people are more cautious about keeping their privacy, it will prove to be more difficult to collect data.

PROVIDING AN EFFECTIVE MECHANISM TO PERFORM SOCIAL INFLUENCE ANALYSIS

The current models for social influence evaluation and the influence maximization problem are simplified, without considering features such as age of users, spatial information of users, and the evolution of relationships among friend nodes over time. In addition, it is simply assumed that a constant spreading rate should be used, despite variations in user propagation rates in practice. In real-world online social network services, such as Twitter and Facebook, each user has a distinct spreading rate on propagating information according to a user's reputation and/or role, such as opinion formers, leaders, and followers. It is a challenging direction to design hybrid approaches that combine the advantages of different algorithms (greedy algorithm, heuristic algorithm, game theory, information entropy, probability the-

ory, etc.) to further improve the efficiency and effectiveness of social influence analysis.

CONCLUSION

In this article, we study social influence analysis in social networking big data from three different perspectives: properties of social influence, relationship between social influence analysis and big data, and architecture of social influence analysis. We also discuss the opportunities for social influence analysis, and provide the challenges for social influence analysis in social networking big data. The current increasing big data from various social networks will provide motivation for further research on social influence analysis, and many new opportunities in many aspects for stimulating industrial and commercial take-up of novel emerging technologies. Along with in-depth analysis of social networking big data, we will face more new challenges in future research. In this article, we contribute to theoretical study on new algorithms and methods of social influence analysis in social networking big data.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China under Grant Nos. 61379041, 61272151, and 61472451; the Key Lab of Guangzhou Mobile Internet Security and Fault Tolerance under Grant No. 2014SY000022; and the High Level Talents Program of Higher Education in Guangdong Province under Funding Support Number 2016ZJ01. Guojun Wang is the corresponding author.

REFERENCES

- [1] C. L. Philip Chen and C. Zhang, "Data-Intensive Applications, Challenges, Techniques and Technologies: A Survey on Big Data," *Info. Sciences*, vol. 275, Aug. 2014, pp. 314–47.
- [2] A. Saulwick and K. Trentelman, "Towards a Formal Semantics of Social Influence," *Knowledge-Based Systems*, vol. 71, 2014, pp. 52–60.
- [3] M. Chen, S. Mao, and Y. Liu, "Big Data: A Survey," *Mobile Network Application*, vol. 19, no. 2, Jan. 2014, pp. 171–209.
- [4] G. Wang et al., "Fine-Grained Feature-Based Social Influence Evaluation in Online Social Networks," *IEEE Trans. Parallel and Distrib. Sys.*, vol. 25, no. 9, Sept. 2014, pp. 2286–96.
- [5] D. Cercel and S. Trausan-Matu, "Opinion Propagation in Online Social Networks: A Survey," *Proc. 4th ACM Int'l. Conf. Web Intelligence, Mining and Semantics*, Thessaloniki, Greece, June 2014.
- [6] S. Peng et al., "Containing Smartphone Worm Propagation with an Influence Maximization Algorithm," *Computer Networks*, vol. 74, Sept. 2014, pp. 103–13.
- [7] R. Tinati et al., "Big Data: Methodological Challenges and Approaches for Sociological Analysis," *Sociology*, vol. 48, no. 1, Feb. 2014, pp. 23–39.
- [8] G. Song et al., "Influence Maximization on Large-Scale Mobile Social Network: A Divide-and-Conquer Method," *IEEE Trans. Parallel and Distrib. Sys.*, vol. 26, no. 5, May 2015, pp. 1379–92.
- [9] S. Zhao and Ting Liu, "Causality and Its Applications in Social Media: A Survey," *J. Software*, vol. 25, no. 12, Dec. 2014 (in Chinese), pp. 2733–52.
- [10] S. He et al., "Identifying Peer Influence in Online Social Networks Using Transfer Entropy," *Proc. Pacific Asia Workshop Intelligence and Security Informatics*, Beijing, China, Aug. 2013, pp. 47–61.
- [11] D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the Spread of Influence in a Social Network," *Proc. 9th ACM SIGKDD Int'l. Conf. Knowledge Discovery and Data Mining*, Washington, 2003, pp. 137–46.
- [12] J. Sun et al., "Leverage RAF to Find Domain Experts on Research Social Networks Services: A Big Data Analytics Methodology with MapReduce Framework," *Int'l. J. Production Economics*, vol. 165, July 2015, pp. 185–93.
- [13] Y. Zhou and L. Liu, "Social Influence Based Clustering of Heterogeneous Information Networks," *Proc. 19th ACM SIGKDD Int'l. Conf. Knowledge Discovery and Data Mining*, Aug. 2013, Chicago, IL, pp. 338–46.

-
- [14] F. Bonchi *et al.*, “Social Network Analysis and Mining for Business Applications,” *ACM Trans. Intelligent Systems and Technology*, vol. 2, no. 3, Apr. 2011, pp. 1–37.

BIOGRAPHIES

SANCHENG PENG received his Ph.D. degree in computer science from Central South University, China, in 2010. Currently, he is a professor at Guangdong University of Foreign Studies, China. He was a research associate at the City University of Hong Kong from 2008 to 2009. His research interests include network and information security, trusted computing, social networks, and mobile computing. He is a member of CCF and ACM.

GUOJUN WANG [M] received his B.Sc. in geophysics, M.Sc. in computer science, and Ph.D. in computer science from Central South University. He is a professor with Guangzhou University.

He was a professor at Central Sount University, China, an adjunct professor at Temple University, a visiting scholar at Florida Atlantic University, a visiting researcher at the University of Aizu, Japan, and a research fellow at Hong Kong Polytechnic University. His research interests include network and information security, the Internet of Things, and cloud computing. He is a distinguished member of CCF, and a member of ACM and IEICE.

DONGQING XIE received his B.Sc. in applied mathematics and M.Sc. in computer software from Xidian University, China; and received his Ph.D. in applied mathematics, from Hunan University, China. He is the Dean and a professor at the School of Computer and Education Software of Guangzhou University. He has been a visiting scholar at Nipissing University, Canada. His research interests include information security and cryptography. He is a member of CCF.