



**INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH  
TECHNOLOGY**

**WELL-DEFINED APPROACHES FOR IDENTIFYING INFLUENTIAL NODES IN A  
SOCIAL NETWORK**

**H.Jayamangala\*, Dr. Ananthi Sheshasaayee**

\* Research Scholar, Periyar University, Salem

Associate Prof. & Head, PG & Research Dept. of Computer Science,  
Quaid-e-Millath Government College for Women (Autonomous), Chennai

---

**ABSTRACT**

Social networks often serve as a medium for spreading information, ideas or innovations. An individual's decision on buying a product or innovation will be highly dependent on the choices made by the individual's peers or neighbors in the social network. Social networks contain graphs of individuals and their relationships, collaborations, or advice seeking relationships. The member or a node that has a large number of contacts with other nodes is referred as an information hub or influential node. Although members have numerous connections, only a fraction of those so called friends may actually influence others. The influential node plays a vital role in diffusing information, including propaganda, ideology, and gossip in a social network. This paper focuses on approaches for finding influential nodes in social network.

**KEYWORDS:** Social network, Influential nodes, Information hub, Web 2.0

---

**INTRODUCTION**

The emergence of Web 2.0 had dramatically changed the growth and usage of the Internet and enabled an interactive and collaborative user experience to advance Internet services dramatically. The popularity and tremendous growth of online social networks (OSNs) has attracted researchers to focus on modeling and properties. Social networks such as Google and Face book involve certain kinds of relationships among either individuals or data. In social networks, communities come together, attract new members, and evolve over a period of time.

Social networking becomes a massive communication medium in the recent years that connects millions of geographically distinct people. People share their thoughts, and information through online social networking sites like Face book [1], LinkedIn [2] and Twitter [3]. Social networks have a wide variety of information in the fields of entertainments, technology, business, sociology and politics. On-line groups are popular in social networking sites and people tend to establish connectivity with other people who have similar features [4]. Information hub has the large number of connections, among other users, which can easily diffuse the information in a social network.

One of the challenges in a social network is finding the information hub. In a social network influence is defined as, the power or capacity of a person or things in causing an effect in indirect or intangible ways. The notion of social influence among people is applied in several applications, namely marketing, advertisement and recommendations. The social influence has three categories such as compliance, identification and internalization. Compliance is an agreement among people by keeping the difference of their opinions private. The identification is the second category in which people influenced by someone whom people likes, respects and follows. The final category is internalization in which people accept a behavior and makes the deal publicly and privately. The influence characteristics depends on several factors such as charisma, reputation, social personality and psychological manipulation abilities.

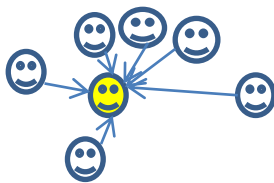
Influence has two forms: positive influence and negative influence. Positive influence is in the form of leadership, sales, marketing and promotion of views. Negative influence is in the form of rumors, gossips, bullying and cussing.

The potency of social influence mainly depends on specific factors such as the strength of relationships between people, the future, the size and popularity of information on the community, distance between

users, and characteristics of networks and individuals [5]. Organizations started to use this information diffusion feature in marketing to improve their product promotion through the concept of word-of-mouth [6][7]. For instance, if a person buys a product he would influence his friends, and they would influence their friends and finally the product would reach many individuals and groups. Small budget companies follow a viral marketing approach that gives free samples to the influential users who strongly scatter information in a network. Many of the organizations exploit word-of-mouth and viral marketing method property to promote their products, innovations, and to diffuse the information in a social network. Information hub is an important tool that effectively diffuses product information, decisions about adoption of a product and other information from one to many nodes or individuals [8].

### CHARACTERISTICS OF INFORMATION HUB IN SOCIAL NETWORKS

In a social network, node indicates individuals and edge indicates interactions among the nodes. The structural properties of the social network were analyzed in [9]. The degree of a node is used to find the hub. A node degree represents the number of connections among other nodes. A node with a high degree is denoted as information hub. Organizations initially select and spread the information to the related information hubs. Information hub is mainly used in marketing and also used in various areas such as public opinion, healthcare, communication, education, agriculture, and epidemiology.



*Figure: Information hub*

The above figure is an example of Information hub which is connected by many members. Hubs are in the center of the network and other members are influenced by the hub. The influential node can be characterized based on the following three factors such as the personification of absolute values, the competence and the strategic social location.

There are three factors which decide influential nodes. The first factor represents a person who has abilities such as communication, successfully convinces people makes him an influential node. The second

factor defines an influential node is an expert in the distinct field of knowledge. The final factor describes the position of an influential within a society.

### THE SIGNIFICANCE OF THE INFORMATION HUB IN SOCIAL NETWORKS

Hubs adopt the product or information so earlier, which increases the speed of adoption in a social network

- Hubs increase not only the adoption speed, but also a market size.
- Influential hubs are convinced that can make another node to accept or adopt the information.
- Influential hubs are more knowledgeable than the other nodes.
- Influential hubs have a large number of connections (social ties) among others.
- Influential hub diffuses the right information to the right people at the right time

### LITERATURE SURVEY

This section discusses several approaches for discovering influential nodes in a social network.

### APPROACHES FOR FINDING INFLUENTIAL NODES

Solving the problem of finding top-k nodes in a social network is discussed in [10]. It is based on a novel and intuitive algorithm called Shapley value that uncovers the most influential nodes approximately [10]. The top-k problem is modeled as a game and the Shapley value gives the marginal contribution to make the coalition dynamics. If the node has higher contribution value and Shapley value, the node is considered to be more important than other nodes. The significance of the algorithm is that it is a generalized approach and the underlying objective function has not used the sub-modularity. Hubs play a vital role in the process of information adoption and diffusion. There are two types of hubs identified in a large network with adoptions, innovative and follower. This process examines the role of hub in influencing the adoption speed and market size [8]. On examining the adoption timing, the relation of hub adoption to market size predicts the future success of a new product at the earlier. Testing the influence of hubs on the dynamics of adoption on more classic products is a significant process.

Community based Greedy algorithm (CGA) [11] extracts the top-K influential nodes in a mobile social network using the structural property of the community. CGA includes two algorithms which

detect the communities among the social network through information diffusion, and selects a community to determine the influential nodes using a dynamic programming. CGA is faster than the existing greedy algorithm. CGA first detects communities and initializes the selected influential nodes to diffuse information. The community detection algorithm follows partition and combination to detect the communities. Moreover, CGA does not consider user preferences during influence diffusion. Several distributed algorithms are available to identify the information hubs. A novel algorithm identifies the information hub through the access of absolute friendship graph without central entity based on the Kempe-McSherry (KM) algorithm. Principal component centrality (PCC) addresses the privacy challenges and measures the centrality of the friendship graph. PCC requires the explicit knowledge of the network structure. The benefit of KM algorithm is low overhead of message exchange. The algorithm preserves the privacy of users since it does not require any personal information from the user [12]. GAUP is a two stage mining algorithm on the basis of an Independent Cascade (IC) model which finds the most influential nodes for a particular concept according to the end user preferences [13]. In the first stage, GAUP calculates user's preference for every user using Singular value Decomposition (SVD)-based Latent Semantic Indexing (LSI) and uses a collaborative filtering approach to select user preferences over a particular topic. The second stage determines top-k nodes, through combining the traditional greedy algorithms with computed user preferences, and also provides a solution for the influence maximization problem. Multi-task Sparse Linear Influence Model (MSLIM) [14] detects an influential node and simultaneously predict the volume for each contagion in a network. MSLIM is based on the linear influence model. MSLIM has two advantages such as it does not require the network structure and detects different sets of the most influential nodes for different contagions. MSLIM adapts the accelerated gradient method (AGM) framework to provide the solution for the proximal mapping.

The greedy algorithm examines about market strategy to maximize the revenue through influence and exploit strategy. Influence stage offers, free samples to select influential nodes or buyers to spread the influence among people. Similarly, the exploit strategy uses greedy pricing strategy to extract the revenue from other customer through influence diffusion [15]. A novel method ranks nodes to extract the influential nodes in the Independent Cascade (IC) model in a

social network [16]. The nodes are ranked using influence degree based on the observed data on information diffusion over a large degree social network. The diffusion probability of every link in the network must be predefined. It calculates the probabilities using iterative Expectation-Maximization (EM) algorithm.

A fast information propagation problem focuses on quick information diffusion through influential node selection for a time bounded situation in a large social network. In order to achieve the process, initially it selects the minimum size of nodes and these nodes diffuse the information to the whole network. However, the optimization of fast information propagation does not work effectively, when the time needed for information diffusion is 1-hop. Thus, a Latency-bounded Minimum Influential Node Selection Algorithm [17] proposed effectively that achieves the fast information diffusion with a 1-hop time count.

A novel data mining method finds social leaders using a frequent pattern discovery [18]. In social networks, log file stores the details of a user's action. A Database combines social graph and stored user actions. Social leaders can extract from the combined value through applying an algorithm. The algorithm depends on the update and propagates functions. If a node of a particular window disappears as a result of movement of the window, it requires an update on the situation of the existing nodes. If a node freshly arrive the window, it propagates the state from existing node to the fresh node. The framework in [5] provides a strategy to identify influential users in large social networks based on the activity level of users called Bayesian shrinkage approach. It finds influential users, using the longitudinal records of users. The Bayesian shrinkage approach is implemented in a Poisson regression model that shrinks influence estimates across friends.

## FUTURE DIRECTIONS

This survey presents the characteristics, importance of information hub and approaches for finding information hub in social networks. The future direction of the information diffusion will focus on several solutions for influence maximization. Identification of the influential nodes successfully in several personalized areas without the prior knowledge of the network structure is a challenging process. Another direction of the future work is to

generate a personal recommender system effectively for the most influential node in a social network.

## CONCLUSION

The online social websites act as an important platform for effective viral marketing practice. In order to promote a company's sales quantity, it is important to find out the node (individual) in the network with high influence. This information hub considerably leads to a much larger spread of the innovation than the traditional view of marketing to individuals. Information hub has the large number of connections, among other users, which can easily diffuse the information in a social network. Thus, the significance of finding the influential node and its applications will grow rapidly in the coming days. Therefore, the need for efficient methods to identify influential nodes has great importance in the near future.

## REFERENCES

1. [www.facebook.com](http://www.facebook.com)
2. [www.linkedin.com](http://www.linkedin.com)
3. [www.twitter.com](http://www.twitter.com)
4. Backstrom, L., Huttenlocher, D., Kleinberg, J., Lan, X., "Group formation in large social networks: membership, growth, and evolution", Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining, pp 44-54, 2006.
5. Trusov, Michael, Anand Bodapati and Randolph E. Bucklin (2010), "Determining Influential Users in Internet Social Networks," Journal of Marketing Research, Vol. 47, No 4, pp 643-658, 2006.
6. J. J. Brown and P. H. Reingen, " Social ties and Word-of-Mouth referral behavior", The Journal of Consumer Research, Vol 14, No 3, pp 350-362, 1987.
7. I. R. Misner. The World's best known marketing secret: Building your business with word-of-mouth marketing. Bard Press, 2nd edition, 1999.
8. Jacob Goldenberg, Sangman Han, Donald R. Lehmann, & Jae Weon Hong," The Role of Hubs in the Adoption Process" Journal of Marketing Vol. 73, 1-13, March 2009.
9. A Mislove, M Marcon, KP Gummedi, P Druschel, B Bhattacharjee, "Measurement and Analysis of Online Social Networks", The 7th ACM SIGCOMM Conference on Internet Measurement, pp 29-42, 2007.
10. N. Rama Suri, Y. Narahari "Determining the Top-k Nodes in Social Networks using the Shapley Value", Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems - Vol 3, pp 1509-1512, 2008.
11. Yu Wang, Gao Cong, Guojie Song, Kunqing Xie," Community-based Greedy Algorithm for Mining Top-K Influential Nodes in Mobile Social Networks", Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining, pp 1039-1048, 2010.
12. Muhammad U. Ilyas, M. Zubair Shafiq, Alex X. Liu," A Distributed Algorithm for Identifying Information Hubs in Social Networks", IEEE Journal on Selected Areas in Communications, Vol 31 , No 9, pp 629-640, 2013.
13. Yunlong Zhang, Jingyu Zhou, and Jia Cheng," Preference-based Top-K Influential Nodes Mining in Social Networks", Proceedings of the 2011 IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications, pp 1512-1518, 2011.
14. Yingze Wang and Guang Xiang and Shi-Kuo Chang," Sparse Multi-task Learning for Detecting Influential Nodes in an Implicit Diffusion Network", Proceedings of 27th AAAI Conference on Artificial Intelligence, 2013.
15. Jason Hartline, Vahab S. Mirrokni, Mukund Sundararajan, "Optimal Marketing Strategies over Social Networks", Proceedings of the 17th international conference on World Wide Web, pp 189-198, 2008.
16. Masahiro Kimura, Kazumi Saito, Ryohei Nakano, and Hiroshi Motoda, "Finding Influential Nodes in a Social Network from Information Diffusion Data", Proceedings of the International Workshop on Social Computing and Behavioral Modeling, pp. 138-145, 2009.
17. Feng Zou, Zhao Zhang, and Weili Wu, "Latency-Bounded Minimum Influential Node Selection in Social Networks", Proceedings of the 4th International Conference on Wireless Algorithms, Systems, and Applications, pp 519 - 526, 2009.
18. Amit Goyal, Francesco Bonchi , Laks V. S. Lakshmanan," Discovering Leaders from Community Actions", Proceedings of the

17th ACM conference on Information and  
knowledge management, pp 499-508, 2000