Journal of Computer and Communications, 2023, 11, 58-73
https://www.scirp.org/journal/jcc
ISSN Online: 2327-5227
ISSN Print: 2327-5219

Identification of Influential Users in Online Social Network: A Brief Overview

Mahmuda Ferdous[1], Md. Musfique Anwar[2]

[1]Department of Computer Science and Engineering, University of Development Alternative, Dhaka, Bangladesh
[2]Department of Computer Science and Engineering, Jahangirnagar University, Dhaka, Bangladesh
Email: ferdous_mahmuda@yahoo.com, manwar@juniv.edu

How to cite this paper: Ferdous, M. and Anwar, M.M. (2023) Identification of Influential Users in Online Social Network: A Brief Overview. *Journal of Computer and Communications*, **11**, 58-73.
https://doi.org/10.4236/jcc.2023.117005

Received: June 1, 2023
Accepted: July 25, 2023
Published: July 28, 2023

Abstract

Information networks where users join a network, publish their own content, and create links to other users are called Online Social Networks (OSNs). Nowadays, OSNs have become one of the major platforms to promote both new and viral applications as well as disseminate information. Social network analysis is the study of these information networks that leads to uncovering patterns of interaction among the entities. In this regard, finding influential users in OSNs is very important as they play a key role in the success above phenomena. Various approaches exist to detect influential users in OSNs, starting from simply counting the immediate neighbors to more complex machine-learning and message-passing techniques. In this paper, we review the recent existing research works that focused on identifying influential users in OSNs.

Keywords

Online Social Network, Trending Topics, Social Influence, Influential User

1. Introduction

Since its creation, the Internet has become a major source of news. The Internet has spawned many information-sharing networks, the most well-known of which is the World Wide Web. Recently, a new category of information networks known as "Online Social Networks (OSNs)" has exploded in popularity and now rivals the traditional Web in terms of usage. The advent of Online Social Networks (OSN) has been one of the most exciting phenomena in the past decade partially due to the increasing proliferation and affordability of Internet-enabled devices, such as personal computers, mobile devices, smartphones, tablets, etc. This is evidenced by the immense popularity of many online social networks, such as

Twitter[1], Facebook[2], LinkedIn[3], Flickr[4], Google+[5], etc.

Unlike the traditional Web, which is largely organized by content, OSN embodies a social structure consisting of a set of users and the ties between them. Users can join a network and create links with other users in the network. This basic user-to-user link structure facilitates online interaction among people in which they produce, share, and exchange user-centric data in a wide variety of scenarios (such as online gaming, instant messaging, and content sharing) in virtual communities and networks. While OSN is mostly used for everyday chatter, it is also used to share news and other important information [1] [2]. Now, many people consider OSN as their source of news [3] [4], this is especially true for breaking news, where people crave rapid updates on developing events in real time. Kwak *et al.* [4] have shown that over 85% of all trending topics[6] on Twitter are headline or persistent news. Mobile technologies offer easy access everywhere and events are reported while happening [5]. The ubiquity, accessibility, speed and ease of use of OSN have made them invaluable sources of first-hand information. Therefore, we can observe a huge amount of data being produced at incredibly rapid rates while information becomes faster and outdated.

Social Network Analysis (SNA) has become a rapidly emerging research discipline in the last decade because of the social and informative properties of the social network. The methods and techniques of SNA involve a variety of areas, including mathematics, statistics, and computer science [6]. Due to its relevance to various processes taking place in society, SNA finds significant applications in different domains, such as sociology, biology, communication, geography, social computing, and business [7] [8] [9]. SNA concentrates on techniques to analyze the relationships and information flows, between nodes/interacting units (people, groups, organizations, etc) in OSN, and produce formal models that facilitate understanding of the structure of a social network.

Nowadays, social network data is everywhere due to its roles as both a media and communication channel. For example, there is data about user-to-user friendship, chat, affiliation, co-author, movie and music networks, etc. Again, data collection is becoming very easy using the API and WebPages provided by those OSN services. The enormous growth of social networking sites and user engagement in these platforms has attracted significant interest from all walks of life. For-profit businesses are tapping into social networks as both a rich source of information and a business execution platform for product design and innovation, relations management and marketing between consumers and stakeholders. For politicians and governments, social networks represent the ideal information base to gauge public opinion on policies as well as to build

[1]http://www.twitter.com/.
[2]http://www.facebook.com/.
[3]http://www.linkedin.com/.
[4]https://www.flickr.com/.
[5]https://plus.google.com/.
[6]Trending topics are those topics being discussed more than others on Twitter.

community support for candidates running for public office.

The analysis of OSNs data sets is turning out to be non-trivial, since on the one hand, they are typically large-scale and thus it is difficult to directly apply the traditional analytics methods to them; on the other hand, some analytic methods that are borrowed from the other domains fail to accurately measure the characteristics and properties of OSN.

Each social networking site includes different gatherings where users can impart their insight, spread data, or convey it to others. This has created a new way to copy genuine human interaction. For example, Twitter has become one of the biggest and very well-known microblogging sites on the internet. Twitter allows registered users to post and receive short messages of up to 140 characters. This feature that allows Twitter users to publish short messages, in a faster and summarized way, makes it the preferred tool for the quick dissemination of information over the web. These messages are called *tweets* and they can be posted via the Twitter website, short messaging services or third-party applications. Importantly, a large fraction of the tweets are posted from mobile devices and services, such as Short Message Service (SMS) messages. A user's messages are displayed as a *stream* on the user's Twitter home page.

As the social networking site (like Twitter) coordinates a massive number of users and the most significant part of the cases is that the users send the *following* request to others who have comparative similar topical interests with the user—the idea is known as homophily [10]. This social platform can be used as a robust marketing platform undoubtedly. The efficiency of online social interactions draws the attention of researchers to discover influential users. Finding influential users can help by highlighting the products in viral marketing or conveying opinions on different social topics among their followers. For example, *Sam*, a sports enthusiast, posts a lot about trending games, and all his posts are paid close attention by a large number of people. A sports company can easily target *Sam* to collaborate to achieve their marketing objectives. This is also called influencer marketing. There are significant impacts of influential users in real life. Some of them are:

- Remarkable impact in viral marketing and social decisions, like elections, donations, etc.
- Anticipating target customers and enhancing customer engagement.
- Analyzing the trends and behavior of people towards any event.
- Tracking down the collective identity and role models in social movements.

In this paper, we review the existing research works to find influential users in OSNs. The rest of the paper is categorized as follows: Section 2 includes the relevant research works in this field with a brief overview. Section 3 covers the common approaches/methodologies in recent times to find influential users in OSNs. Section 4 shows comparisons over some existing research methods considering different attributes and performance strategies, and Section 5 concludes our review work with the direction of future research scope.

2. Related Work

The social networks are becoming complex, and directly reaching out the users is becoming more challenging. As a result, finding influential users has become a key issue in viral marketing. An online relationship refers to user's virtual social relationships in social networks, such as the *following* relationships on Twitter. They can be typically expressed as directed edges between each pair of users in social graph.

Figure 1 shows a directed social network comprised of four nodes with their related messages. This representation reveals that, for example, the user named "u_1" is exposed to the content produced by "u_2" and "u_3". It also indicates that none of the three other nodes are exposed to the information shared by "u_4".

Every user-generated content contains one or more topics. Usually, the contents published by the users of OSNs are viewed as a stream of messages. Figure 1 represents the stream produced by the members of the network depicted in the previous example. That stream can be viewed as a sequence of users' activities.

Social Influence: A social phenomenon that individuals can undergo or exert, also called imitation, translating the fact that actions of a user can induce his connections to behave in a similar way. Influence appears explicitly when someone shares or like content posted by someone else for example.

Existing methods of finding influential users in OSNs can be generally grouped into two categories: structural and hybrid methods. In the following, we review these two categories.

2.1. Structural Methods

Social connections among the users can be used to measure the popularity or influence of users, in which high-degree node (user) is assumed as the authority for the largest information dissemination [11]. These social connections can be either directed (for example, follower/followee) or undirected (friendship relations in Facebook). In-degree centrality refers to the number of edges that connect to the node, whereas out-degree centrality indicates the number of edges that originate from the node. In directed networks, in-degree centrality usually

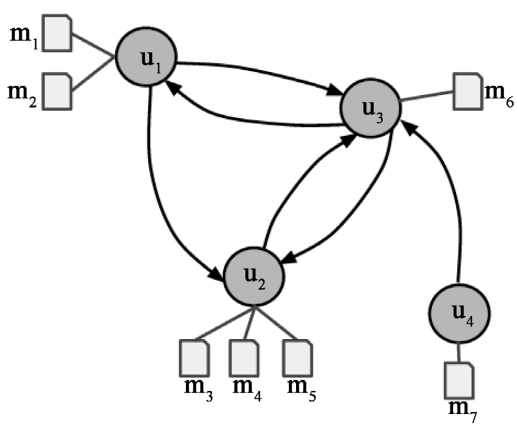

Figure 1. The stream of messages produced by the users of the social network.

refers to the popularity of a user, whereas out-degree centrality typically indicates the sociality of a user [12].

Again, *betweenness centrality* of a user is calculated by the counts of the shortest paths that pass through that user to identify influential users in OSNs. For example, Catanese *et al.* [13] have applied betweenness centrality to Facebook social graph to identify the central nodes of the network. Similarly, Katz [14] determined centrality's influence on the node by all network links that pass through the node.

The TURank algorithm [15] studies the relationship among users along with the users' posts by taking into account the relationship graph network of user-to-tweets. The TwitterRank algorithm [16] uses the topics discussed on Twitter along with the network structure to rank user influence on Twitter.

- Freeman [17] considered nodes' (users') degree centrality, *i.e.* the more neighbors a node (user) has, the higher is its influentiality.
- Kitsak *et al.* [18] chose k-shell centrality which is a k-shell index that is assigned to each node indicating its distance to the network core. Nodes with higher k-shells are considered more influential as they are closer to the graph core.
- Sheikhahmadi *et al.* [19] proposed a method that considered user's connections in different shells along with the k-shell and degree measures for presentation of a hybrid measure for determining spreading capability.
- Chen *et al.* [20] proposed an improved version of the degree centrality measure. The algorithm is iterated k times to select k nodes, and the node with the highest degree is selected and added to the seed set in each iteration. The edges between the selected node and the other network nodes are disregarded in specification of spreading capability of the nodes.
- In [21], authors proposed a method where nodes are first colored such that nodes of the same color will have distance which is higher than a certain threshold value. Then, the nodes are grouped and ranked based on their color and their degree. Finally, the top-k nodes with highest degree within the group are selected as the most influential nodes.

2.2. Hybrid Methods

Awan *et al.* [22] have predicted the stock market using big data retrieved from social media like Yahoo!, daily newspaper, and Twitter. Even policing protests in the United Kingdom are analyzed by social media data [23]. Additionally, cyber risk management [24], mental health condition [25], suicide rate and causes [26], box office's profit [27], etc. are predicted through social media analysis. The prominent Influencers' impact on consumer behavior is evaluated in the work of Pick [28]. The effectiveness of sponsors and influential users' in multiple domestic and business sectors are analyzed in the work of Feng *et al.* [29]. The study of Anuar *et al.* [30] has found out the cause of being influenced by Instagram influencers in regards to purchasing intention of fashion items.

Earlier approaches focused on the immediate neighbors (for example, counting the neighbors) for detecting influential. One of the first studies that attempted to find the parameters of this approach was taken by Zhang *et al.* [31]. They have considered users' retweet behavior patterns to investigate how friends in one's ego network influence retweet behaviors. In this model, the designs incorporate the social influence locality into a factor graph model, further leveraging the network-based correlation. Weng *et al.* [16] have suggested a measure named TwitterRank based on the idea of PageRank to compute users' topical influence in Twitter. This approach is based on the topic query set and it shows the relevance of link structure and the similar interest of users. Al-garadi *et al.* [32] have calculated the users' interactions and modeled the social graph using the weighted *k*-core decomposition method to identify the influential spreaders in OSNs. To identify top-*k* significant users in social networks, Alshahrani *et al.* [33] have proposed an efficient algorithm based on centrality measures. Moreover, Zareie *et al.* [34] have proposed a method to select the influential users based on the interest value of friends' interests and connected neighborhood. A new approach named Temporal Topic Influence (TTI) has been proposed by Wang *et al.* [35] states that analytical applications in online social networks can be generalized as the influence evaluation problem, which targets at finding most influential users. This model is dependent on time interval, content, and structure-aware. Othman *et al.* [36] have investigated the effect of topic familiarity on listening comprehension and how far certain aspects of the language would likely be influenced by topic familiarity. The UIRank algorithm is based on the commitment of the user's tweet and the attributes of information dispersal in the microblog networks. It computes user influence score iteratively by user follower graph [37]. Most of them ignore the time factors in their work.

To discover highly reliable domain-based influencers at different time intervals, Abu-Salih *et al.* have suggested a framework with the help of semantic analysis and machine learning modules [38]. Again, there is an on-Demand Influencer Discovery (DID) model, which employs an iterative learning process incorporating the language attention network as a subject filter, proposed by Zang *et al.* that can identify influential users on any subject regardless of its demand on social media [39]. Their influence convolution network is built on user interaction. But they did not suggest any rank for their influential users. The research of Mittal *et al.* has discovered and ranked significant users (topic wise) [40]. Their proposed Aggregation Consensus Rank Algorithm (ACRA) is applied on time intervals to generate top-ranked influential users' lists using different Twitter metrics. They analyze the connection between users and graph database to find this significant user. In a framework named Personalized PageRank that also identifies influential topical users based on both information gathered from the network and the data retrieved from user actions [41]. Additionally, fake influencers can be a threat to marketing and advertising. A trust-based method for identifying these inorganic users is proposed by Dewan [42] using decision tree and machine learning.

Among the recent research, Li *et al.* [43] work on sensitive influence maximization on the different interesting topics of different users. Their proposed algorithm is based on graph pruning and a three-stage heuristic optimization strategy. Mandal *et al.* proposed Social Promoter Score (SPS)-based recommendation [44]. Kumar *et al.* [45] found Top-k influential nodes in a community using label propagation. They claimed their work using several real-life data. Another approach of influence maximization is proposed by Li *et al.* [46]. Their framework is based on a meta-heuristic search algorithm. In a social network, Shi *et al.* [47] proposed a community detection algorithm established on Quasi-Laplacian centrality peaks clustering.

But, most of the existing approaches overlooked the combination of analyzing trending topics and the temporal factor, which significantly affects the ranking of the influential users. Our current proposed method is the extension of Topical Influential Users Detection (TIUD) algorithm [48]. This is also defined as finding significant users for a set of trending topics and listing the top significant users at different specific time intervals considering familiar neighbors.

3. Overview of Research Method

3.1. Problem Formulation

We first give some fundamental concepts related to the task of identifying influential users in OSNs.

Social Graph: An attributed social graph manifests as $G=(U,E,A)$, where U symbolizes the set of social users or twitterers (nodes), E means the group of links (edges) in the users and $A=\{T_1,T_2,\cdots,T_n\}$ is the set of topics contemplated by the social users in G [49] [50] [51].

Topic: Any specific keyword or a set of related words which illustrates equal thought can be assessed as the topic [52] [53] [54]. For instance, when *health* is a topic, words linked to health are like doctors, hospital, pandemic, corona, etc.

Trending Topic: A trending topic is a concern that meets an inundation of popularity, often advancing around widespread contemporaneous phenomenons. For example, celebutantes' pronouncements, breaking news, social affairs, etc., for a limited range of time [55] [56] [57] [58].

Activity: Social users execute actions (e.g. posting tweets on Twitter) at various time points. This activity is recorded as an activity tuple (u_i,Ψ_{u_i},t_n) where Ψ_{u_i} exemplifies the set of the high-ranked trending topics which is covered by the working that is conducted by u_i at time t_n [59] [60] [61].

Query: An input query $Q=\{\Psi_q\}$ assimilates a set of query topics Ψ_q.

3.2. Topic Modelling

Social media like Twitter contains more compact messages so generally one tweet is the reflection of one topic. Here, users normally tweet by using hashtags (for example, #Obama, #Ronaldo, etc.). Topic modeling is an unsupervised learning that generates information and analyzes words from documents by linking words with the same features and differentiates across the uses of words with various meanings. To retrieve the textual content from a tweet, Twitter LDA [62] (T-LDA), an effective extension of LDA is used here for topic distillation. Twitter-LDA is better in topic semantic coherence by presuming the ratio between topic and background words is indifferent for each user's tweets. The graphical representation of T-LDA is shown in **Figure 2**. The formulation of T-LDA [63] is given below:

- Every individual user's topical interest ϕ_i is represented by a distribution over N topics.
- Each word is implied by topic N is analyzed from a background word distribution represented by θ_{bw} or topic word distribution θ_{kw}.
- If The latent value $y=0$, it verifies that the word is from background word distribution θ_{bw} and if $y=1$, it is from topic word distribution θ_{kw}.
- y is altered based on the ratio of background words and topic word denoted by π. π is the common factor where the rate of θ_{kw} and θ_{bw} is same.
- DT, a $D \times T$ matrix, where D is the number of Twitter users and T is the number of topics. DT_{ij} contains the number of times a word in twitterer s_i's tweets has been assigned to topic t_j [16].

3.3. Influential User Detection Approach

Recent methodologies usually apply PageRank-like algorithm [64] to find influential users in an online social attributed graph G for a given query Q. The framework has the following steps (depicted in **Figure 3**):

- Identify the set of topics using tweets of different users over different time periods.

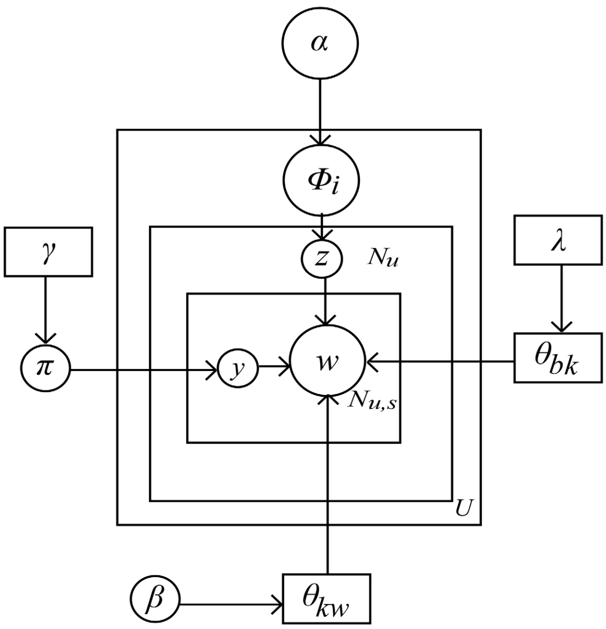

Figure 2. Graphical depiction of Twitter LDA model.

- Modify the original social graph G by constructing topic specific relation among users.
- Apply TwitterRank on the modified social graph in order to detect the influential user.

3.4. Topic Distribution

One needs to apply Twitter LDA [62] (T-LDA) in order to extract the topics discussed by the users in Twitter. Table 1 represents sample word-topic distribution

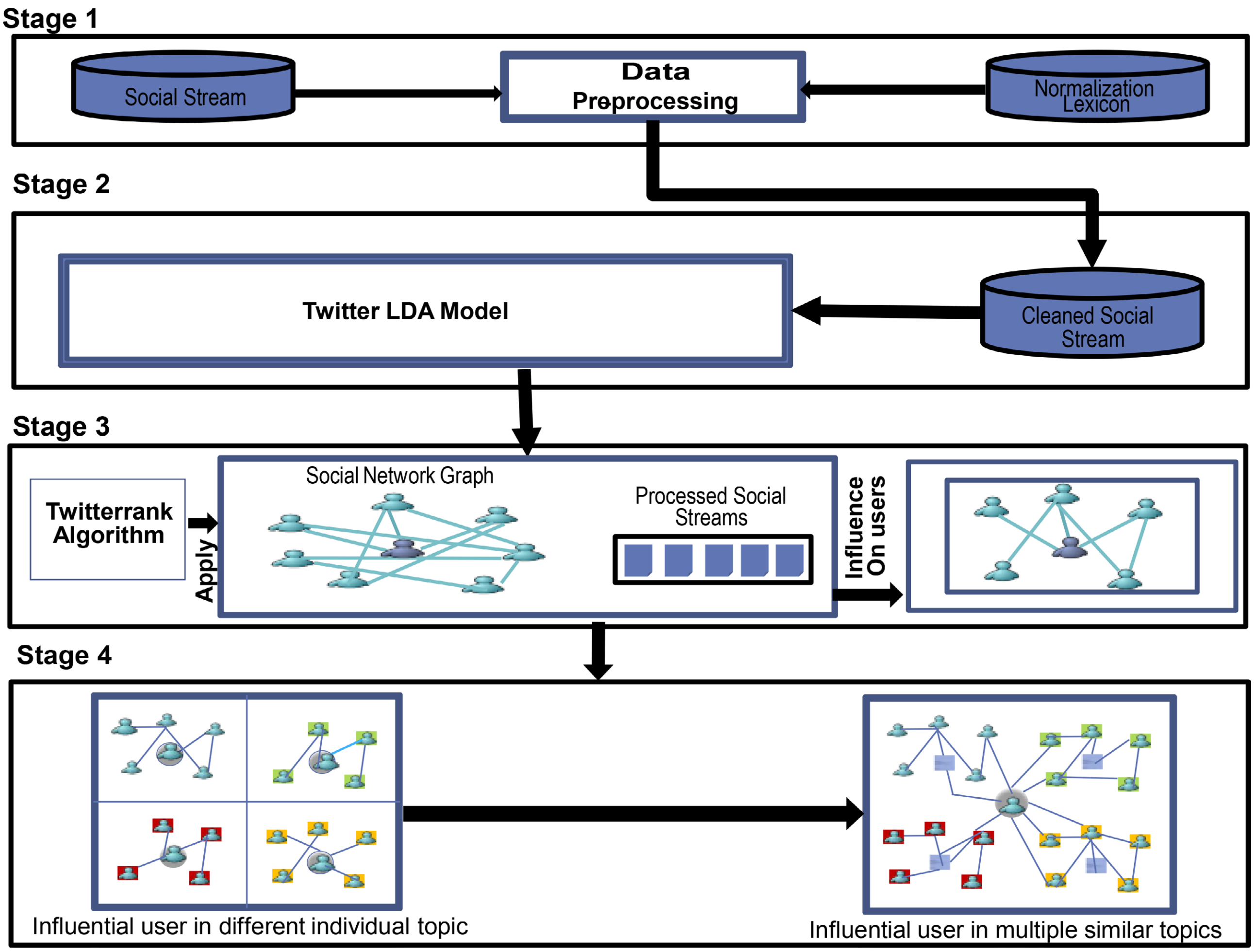

Figure 3. Workflow for identifying influential users in OSNs [56].

Table 1. Sample word topic modeling in Twitter LDA.

Topic I	Topic II	Topic III	Topic IV
1. Sandwich	1. YouTube	1. Flickr	1. Ronaldo
2. Eggs	2. Twitter	2. iPad	2. Football
3. Coke	3. Facebook	3. Watch	3. FIFA
4. Puffer Fish	4. Fun	4. Tweet-bot	4. Cup
5. Butter	5. Instagram	5. Desktop	5. Badminton

after applying Twitter-LDA in our experimental dataset. Topic I, Topic II, Topic III, and Topic IV denote the topics about *food*, *entertainment*, *technology* and *sports* respectively [61].

4. Comparison

In this section, we present a very brief comparison of some existing methods in Table 2 by considering different characteristics, such as identification algorithms, performance evaluation (artificial/manual/direct comparison) and attributes.

Table 2. Comparison of the identification of influential users in OSNs [65].

Research Method	Technique/Alg			Attribute		Evaluation		
	Degree Centrality	Betweenness Centrality	Page Rank-like	Network	Content	Artificial Model	Manual Annotation	Direct Comparison
Cha *et al.* [66]	√	-	-	√	√	-	-	√
Kim *et al.* [67]	√	-	-	√	√	√	-	-
Catanese *et al.* [13]	√	√	-	√	-	-	-	√
Li *et al.* [68]	-	-	√	√	-	√	-	-
Cossu *et al.* [69]	√	√	√	√	√	-	√	-
Sun *et al.* [70]	-	-	-	√	√	-	-	√
Zhuang *et al.* [71]	-	-	-	√	√	√	-	-
Tan *et al.* [72]	-	-	-	√	√	√	-	-
Weng *et al.* [16]		-	√	√	√	-	-	√
Gomasta *et al.* [61]		-	√	√	√	-	-	√
Xia *et al.* [73]	-	-	-	√	√	√	-	-
Sheikhahmadi *et al.* [19]	-	-	-	√	√	√	-	-

5. Conclusion

In this paper, we briefly review existing research works on finding influential users in OSNs. Earlier methods mostly focused on social connections with less attention to the content generated by social users. Later approaches considered both network and content in order to get topic-oriented influential users. But, the major limitation of those approaches is that they overlook the combination of analyzing trending topics and the temporal factor, which significantly affects the ranking of influential users. Most recent works now pay more attention to dynamic networks to track time-based user activities to rank the most influential users in OSNs. In addition, social users have different degrees of interest in different topics that vary over time and as a result, users' social influences also change over time. Researchers need to focus more on dynamic social graphs where the temporal factor has a great impact on both the social connections and users' activities in order to find temporal influential users on trending topics.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Java, A., Song, X., Finin, T. and Tseng, B. (2007) Why We Twitter: Understanding Microblogging Usage and Communities. *Proceedings of the* 9*th WebKDD and* 1*st SNA-KDD* 2007 *Workshop on Web Mining and Social Network Analysis*, San Jose, 12 August 2007, 56-65. https://doi.org/10.1145/1348549.1348556

[2] Naaman, M., Boase, J. and Lai, C.-H. (2010) Is It Really about Me?: Message Content in Social Awareness Streams. *Proceedings of the* 2010 *ACM Conference on Computer Supported Cooperative Work*, Savannah, 6-10 February 2010, 189-192. https://doi.org/10.1145/1718918.1718953

[3] Stassen, W. (2010) Your News in 140 Characters: Exploring the Role of Social Media in Journalism. *Global Media Journal*: *African Edition*, **4**, 116-131. https://doi.org/10.5789/4-1-15

[4] Kwak, H., Lee, C., Park, H. and Moon, S. (2010) What Is Twitter, a Social Network or a News Media? *Proceedings of the* 19*th International Conference on World Wide Web*, Raleigh, 26-30 April 2010, 591-600. https://doi.org/10.1145/1772690.1772751

[5] Sakaki, T., Okazaki, M. and Matsuo, Y. (2010) Earthquake Shakes Twitter Users: Real-Time Event Detection by Social Sensors. *Proceedings of the* 19*th International Conference on World Wide Web*, Raleigh, 26-30 April 2010, 851-860. https://doi.org/10.1145/1772690.1772777

[6] Wasserman, S. and Faust, K. (1994) Social Network Analysis: Methods and Applications. Vol. 8, Cambridge University Press, Cambridge.

[7] Borgatti, S.P., Mehra, A., Brass, D.J. and Labianca, G. (2009) Network Analysis in the Social Sciences. *Science*, **323**, 892-895. https://doi.org/10.1126/science.1165821

[8] Bonchi, F., Castillo, C., Gionis, A. and Jaimes, A. (2011) Social Network Analysis and Mining for Business Applications. *ACM Transactions on Intelligent Systems and Technology*, **2**, Article No. 22. https://doi.org/10.1145/1961189.1961194

[9] Scott, J. (2017) Social Network Analysis. SAGE Publications, Thousand Oaks. https://doi.org/10.4135/9781529716597

[10] McPherson, M., Smith-Lovin, L. and Cook, J.M. (2001) Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, **27**, 415-444. https://doi.org/10.1146/annurev.soc.27.1.415

[11] Albert, R., Jeong, H. and Barabási, A.-L. (2000) Error and Attack Tolerance of Complex Networks. *Nature*, **406**, 378-382. https://doi.org/10.1038/35019019

[12] Jiang, J., Wilson, C., Wang, X., Sha, W., Huang, P., Dai, Y. and Zhao, B.Y. (2013) Understanding Latent Interactions in Online Social Networks. *ACM Transactions on the Web* (*TWEB*), **7**, Article No. 18. https://doi.org/10.1145/2517040

[13] Catanese, S., De Meo, P., Ferrara, E., Fiumara, G. and Provetti, A. (2012) Extraction and Analysis of Facebook Friendship Relations. In: Abraham, A., Ed., *Computational Social Networks*, Springer, London, 291-324. https://doi.org/10.1007/978-1-4471-4054-2_12

[14] Katz, L. (1953) A New Status Index Derived from Sociometric Analysis. *Psychometrika*, **18**, 39-43. https://doi.org/10.1007/BF02289026

[15] Yamaguchi, Y., Takahashi, T., Amagasa, T. and Kitagawa, H. (2010) TURank: Twitter

User Ranking Based on User-Tweet Graph Analysis. In: Chen, L., Triantafillou, P. and Suel, T., Eds., *WISE 2010: International Conference on Web Information Systems Engineering*, Vol. 6488, Springer, Berlin, 240-253. https://doi.org/10.1007/978-3-642-17616-6_22

[16] Weng, J., Lim, E.-P., Jiang, J. and He, Q. (2010) TwitterRank: Finding Topic-Sensitive Influential Twitterers. *Proceedings of the 3rd ACM International Conference on Web Search and Data Mining*, New York, 4-6 February 2010, 261-270. https://doi.org/10.1145/1718487.1718520

[17] Freeman, L.C. (1978) Centrality in Social Networks Conceptual Clarification. *Social Networks*, **1**, 215-239. https://doi.org/10.1016/0378-8733(78)90021-7

[18] Kitsak, M., *et al*. (2010) Identification of Influential Spreaders in Complex Networks. *Nature Physics*, **6**, 888-893. https://doi.org/10.1038/nphys1746

[19] Sheikhahmadi, A. and Nematbakhsh, M.A. (2017) Identification of Multi-Spreader Users in Social Networks for Viral Marketing. *Journal of Information Science*, **43**, 412-423. https://doi.org/10.1177/0165551516644171

[20] Chen, W., Wang, Y. and Yang, S. (2009) Efficient Influence Maximization in Social Networks. *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Paris, 28 June-1 July 2009, 199-208. https://doi.org/10.1145/1557019.1557047

[21] Guo, L., Lin, J.-H., Guo, Q. and Liu, J.-G. (2016) Identifying Multiple Influential Spreaders in Term of the Distance-Based Coloring. *Physics Letters A*, **380**, 837-842. https://doi.org/10.1016/j.physleta.2015.12.031

[22] Awan, M.J., Rahim, M.S.M., Nobanee, H., Munawar, A., Yasin, A. and Zain, A.M. (2021) Social Media and Stock Market Prediction: A Big Data Approach. *Computers, Materials & Continua*, **67**, 2569-2583. https://doi.org/10.32604/cmc.2021.014253

[23] Dencik, L., Hintz, A. and Carey, Z. (2018) PPrediction, Pre-Emption and Limits to Dissent: Social Media and Big Data Uses for Policing Protests in the United Kingdom. *New Media & Society*, **20**, 1433-1450. https://doi.org/10.1177/1461444817697722

[24] Subroto, A. and Apriyana, A. (2019) Cyber Risk Prediction through Social Media Big Data Analytics and Statistical Machine Learning. *Journal of Big Data*, **6**, Article No. 50. https://doi.org/10.1186/s40537-019-0216-1

[25] Sekulic, I. and Strube, M. (2019) Adapting Deep Learning Methods for Mental Health Prediction on Social Media. *Proceedings of the 5th Workshop on Noisy User-Generated Text* (*W-NUT* 2019), Hong Kong, 4 November 2019, 322–327. https://doi.org/10.18653/v1/D19-5542

[26] Wang, N., Luo, F., Shivtare, Y., Badal, V.D., Subbalakshmi, K., Chandramouli, R. and Lee, E. (2021) Learning Models for Suicide Prediction from Social Media Posts. *Proceedings of the 7th Workshop on Computational Linguistics and Clinical Psychology*: *Improving Access*, Online, 11 June 2021, 87-92. https://doi.org/10.18653/v1/2021.clpsych-1.9

[27] Choudhery, D. and Leung, C.K. (2017) Social Media Mining: Prediction of Box Office Revenue. *Proceedings of the 21st International Database Engineering & Applications Symposium*, Bristol, 12-14 July 2017, 20-29. https://doi.org/10.1145/3105831.3105854

[28] Pick, M. (2020) Psychological Ownership in Social Media Influencer Marketing. *European Business Review*, **33**. https://doi.org/10.1108/EBR-08-2019-0165

[29] Feng, Y., Chen, H. and Kong, Q. (2021) An Expert with Whom I Can Identify: The

Role of Narratives in Influencer Marketing. *International Journal of Advertising*, **40**, 972-993. https://doi.org/10.1080/02650487.2020.1824751

[30] Anuar, N.I.M., Mohamad, S.R., Zulkiffli, W.F.W., Hashim, N.A.A.N., Abdullah, A.R., Rasdi, A.L.M., Hasan, H., Abdullah, T., Deraman, S.N.S., Zainuddin, S.A., *et al*. (2020) Impact of Social Media Influencer on Instagram User Purchase Intention towards the Fashion Products: The Perspectives of Students. *European Journal of Molecular & Clinical Medicine*, **7**, 2589-2598.

[31] Zhang, J., Tang, J., Li, J., Liu, Y. and Xing, C. (2015) Who Influenced You? Predicting Retweet via Social Influence Locality. *ACM Transactions on Knowledge Discovery from Data* (*TKDD*), **9**, Article No. 25. https://doi.org/10.1145/2700398

[32] Al-garadi, M.A., Varathan, K.D. and Ravana, S.D. (2017) Identification of Influential Spreaders in Online Social Networks Using Interaction Weighted K-Core Decomposition Method. *Physica A*: *Statistical Mechanics and Its Applications*, **468**, 278-288. https://doi.org/10.1016/j.physa.2016.11.002

[33] Alshahrani, M., Fuxi, Z., Sameh, A., Mekouar, S. and Huang, S. (2020) Efficient Algorithms Based on Centrality Measures for Identification of Top-K Influential Users in Social Networks. *Information Sciences*, **527**, 88-107. https://doi.org/10.1016/j.ins.2020.03.060

[34] Zareie, A., Sheikhahmadi, A. and Jalili, M. (2019) Identification of Influential Users in Social Networks Based on Users' Interest. *Information Sciences*, **493**, 217-231. https://doi.org/10.1016/j.ins.2019.04.033

[35] Wang, F., Li, J., Jiang, W. and Wang, G. (2017) Temporal Topic-Based Multi-Dimensional Social Influence Evaluation in Online Social Networks. *Wireless Personal Communications*, **95**, 2143-2171. https://doi.org/10.1007/s11277-017-4047-0

[36] Othman, J. and Vanathas, C. (2017) Topic Familiarity and Its Influence on Listening Comprehension. *The English Teacher*, **34**, 19-32.

[37] Zhao, J., Gui, X. and Tian, F. (2017) A New Method of Identifying Influential Users in the Micro-Blog Networks. *IEEE Access*, **5**, 3008-3015. https://doi.org/10.1109/ACCESS.2017.2672680

[38] Abu-Salih, B., Chan, K.Y., Al-Kadi, O., Al-Tawil, M., Wongthongtham, P., Issa, T., Saadeh, H., Al-Hassan, M., Bremie, B. and Albahlal, A. (2020) Time-Aware Domain-Based Social Influence Prediction. *Journal of Big Data*, **7**, Article No. 10. https://doi.org/10.1186/s40537-020-0283-3

[39] Zheng, C., Zhang, Q., Young, S. and Wang, W. (2020) On-Demand Influencer Discovery on Social Media. *Proceedings of the* 29*th ACM International Conference on Information & Knowledge Management*, Online, 19-23 October 2020, 2337-2340. https://doi.org/10.1145/3340531.3412134

[40] Mittal, D., Suthar, P., Patil, M., Pranaya, P., Rana, D.P. and Tidke, B. (2020) Social Network Influencer Rank Recommender Using Diverse Features from Topical Graph. *Procedia Computer Science*, **167**, 1861-1871. https://doi.org/10.1016/j.procs.2020.03.205

[41] Alp, Z.Z. and Öğüdücü, Ş.G. (2018) Identifying Topical Influencers on Twitter Based on User Behavior and Network Topology. *Knowledge-Based Systems*, **141**, 211-221. https://doi.org/10.1016/j.knosys.2017.11.021

[42] Dewan, A. (2021) Detecting Organic Audience Involvement on Social Media Platforms for Better Influencer Marketing and Trust-Based E-Commerce Experience. In: Khanna, A., Gupta, D., Pólkowski, Z., Bhattacharyya, S. and Castillo, O., Eds., *Data Analytics and Management*: *Lecture Notes on Data Engineering and Commu-*

nications Technologies, Vol. 54, Springer, Singapore, 661-673. https://doi.org/10.1007/978-981-15-8335-3_51

[43] Li, Y., Li, R., Xiong, X., Gu, X., Liang, T., Xu, M. and Yuan, Y. (2021) Multi-Topical Authority Sensitive Influence Maximization with Authority Based graph Pruning and Three-Stage Heuristic Optimization. *Applied Intelligence*, **51**, 8432-8450. https://doi.org/10.1007/s10489-021-02213-9

[44] Mandal, S. and Maiti, A. (2021) Deep Collaborative Filtering with Social Promoter Score-Based User-Item Interaction: A New Perspective in Recommendation. *Applied Intelligence*, **51**, 7855-7880. https://doi.org/10.1007/s10489-020-02162-9

[45] Kumar, S., Singhla, L., Jindal, K., Grover, K. and Panda, B. (2021) IM-ELPR: Influence Maximization in Social Networks Using Label Propagation Based Community Structure. *Applied Intelligence*, **51**, 7647-7665. https://doi.org/10.1007/s10489-021-02266-w

[46] Li, H., Zhang, R., Zhao, Z., Liu, X. and Yuan, Y. (2021) Identification of Top-*K* Influential Nodes Based on Discrete Crow Search Algorithm Optimization for Influence Maximization. *Applied Intelligence*, **51**, 7749-7765. https://doi.org/10.1007/s10489-021-02283-9

[47] Shi, T., Ding, S., Xu, X. and Ding, L. (2021) A Community Detection Algorithm Based on Quasi-Laplacian Centrality Peaks Clustering. *Applied Intelligence*, **51**, 7917-7932. https://doi.org/10.1007/s10489-021-02278-6

[48] Dhali, A., Gomasta, S.S., Mohanta, S. and Anwar, M.M. (2020) Identification of Query-Oriented Influential Users in Online Social Platform. 2020 *IEEE Region* 10 *Symposium* (*TENSYMP*), Dhaka, 5-7 June 2020, 973-976. https://doi.org/10.1109/TENSYMP50017.2020.9230644

[49] Aurpa, T.T., Khan, F. and Anwar, M.M. (2020) Discovering and Tracking Query Oriented Topical Clusters in Online Social Networks. 2020 *IEEE Region* 10 *Symposium* (*TENSYMP*), Dhaka, 5-7 June 2020, 1054-1057. https://doi.org/10.1109/TENSYMP50017.2020.9230994

[50] Das, S. and Anwar, M.M. (2019) Discovering Topic-Oriented Highly Interactive Online Communities. *Frontiers in Big Data*, **2**, Article 10. https://doi.org/10.3389/fdata.2019.00010

[51] Anwar, M.M., Liu, C. and Li, J. (2018) Uncovering Attribute-Driven Active Intimate Communities. In: Wang, J., Cong, G., Chen, J. and Qi, J., Eds., *Databases Theory and Applications*. *ADC* 2018: *Australasian Database Conference*, Vol. 10837, Springer, Cham, 109-122. https://doi.org/10.1007/978-3-319-92013-9_9

[52] Anwar, M.M. (2020) Query-Oriented Temporal Active Intimate Community Search. In: Borovica-Gajic, R., Qi, J. and Wang, W., Eds., *Databases Theory and Applications*. *ADC* 2020: *Australasian Database Conference*, Vol. 12008, Springer, Cham, 206-215. https://doi.org/10.1007/978-3-030-39469-1_17

[53] Rahman, M.H., Prama, T.T. and Anwar, M.M. (2021) Modeling Topic Specific Credibility in Twitter Based on Structural and Attribute Properties. In: Abraham, A., Hanne, T., Castillo, O., Gandhi, N., Nogueira Rios, T. and Hong, TP., Eds., *Hybrid Intelligent Systems*. *HIS* 2020: *International Conference on Hybrid Intelligent Systems*, Vol. 1375, Springer, Cham, 580-589. https://doi.org/10.1007/978-3-030-73050-5_57

[54] Setu, S.J., Islam, T., Bhuiyan, M.A.-A. and Anwar, M.M. (2020) TTRank: A Temporal Model to Rank Online Twitter Users. *International Journal of Automation, Artificial Intelligence and Machine Learning*, **1**, 42-53.

[55] Jamil, R., Khan, M.A.A.N. and Anwar, M.M. (2022) Topic Oriented Hate Speech

Detection. In: Abraham, A., *et al*., Eds., *Hybrid Intelligent Systems*. *HIS* 2021: *International Conference on Hybrid Intelligent Systems*, Vol. 420, Springer, Cham, 365-375. https://doi.org/10.1007/978-3-030-96305-7_34

[56] Shahriar, K.T., Islam, M.N., Anwar, M.M. and Sarker, I.H. (2022) COVID-19 Analytics: Towards the Effect of Vaccine Brands through Analyzing Public Sentiment of Tweets. *Informatics in Medicine Unlocked*, **31**, Article ID: 100969. https://doi.org/10.1016/j.imu.2022.100969

[57] Ahmed, M.S., Aurpa, T.T. and Anwar, M.M. (2021) Detecting Sentiment Dynamics and Clusters of Twitter Users for Trending Topics in COVID-19 Pandemic. *PLOS ONE*, **16**, e0253300. https://doi.org/10.1371/journal.pone.0253300

[58] Hossen, M., Afrose, T., Ghosh, A.M. and Anwar, M.M. (2021) Classification of Social Media Users Based on Temporal Behaviors and Interests. In: Sharma, H., Gupta, M.K., Tomar, G.S. and Lipo, W., Eds., *Communication and Intelligent Systems*: *Lecture Notes in Networks and Systems*, Vol. 204, Springer, Singapore, 935-944. https://doi.org/10.1007/978-981-16-1089-9_72

[59] Anwar, M.M., Liu, C. and Li, J. (2019) Discovering and Tracking Query Oriented Active Online Social Groups in Dynamic Information Network. *World Wide Web*, **22**, 1819-1854. https://doi.org/10.1007/s11280-018-0627-5

[60] Das, B.C., Anwar, M.M., Bhuiyan, M.A.A., Sarker, I.H., Alyami, S.A. and Moni, M.A. (2021) Attribute Driven Temporal Active Online Community Search. *IEEE Access*, **9**, 93976-93989. https://doi.org/10.1109/ACCESS.2021.3093368

[61] Gomasta, S.S., Dhali, A., Anwar, M.M. and Sarker, I.H. (2022) Query-Oriented Topical Influential Users Detection for Top-*K* Trending Topics in Twitter. *Applied Intelligence*, **52**, 13415-1343. https://doi.org/10.1007/s10489-022-03582-5

[62] Zhao, W. X., Jiang, J., Weng, J., He, J., Lim, E., Yan, H. and Li, X. (2011). Comparing Twitter and Traditional Media Using Topic Models. In: Clough, P., *et al*., Eds., *Advances in Information Retrieval*. *ECIR* 2011: *European Conference on Information Retrieval*, Vol. 6611, Springer, Berlin, 338-349. https://doi.org/10.1007/978-3-642-20161-5_34

[63] Sasaki, K., Yoshikawa, T. and Furuhashi, T. (2019) Online Topic Model for Twitter Considering Dynamics of User Interests and Topic Trends. *Proceedings of the* 2014 *Conference on Empirical Methods in Natural Language Processing* (*EMNLP*), Doha, 25-29 October 2014, 1977-1985. https://www.aclweb.org/anthology/D14-1212.pdf

[64] Brin, S. and Page, L. (1998) The Anatomy of a Large Scale Hyper Textual Web Search Engine. *Computer Networks and ISDN Systems*, **30**, 107-117. https://doi.org/10.1016/S0169-7552(98)00110-X

[65] Al-Garadi, M.A., Varathan, K.D., Ravana, S.D., Ahmed, E., Mujtaba, G., Khan, M.U.S. and Khan, S.U. (2018) Analysis of Online Social Network Connections for Identification of Influential Users: Survey and Open Research Issues. *ACM Computing Surveys* (*CSUR*), **51**, Article No. 16. https://doi.org/10.1145/3155897

[66] Cha, M., Haddadi, H., Benevenuto, F. and Gummadi, K. (2010) Measuring User Influence in Twitter: The Million Follower Fallacy. *Proceedings of the International AAAI Conference on Web and Social Media*, **4**, 10-17. https://doi.org/10.1609/icwsm.v4i1.14033

[67] Kim, E.S. and Han, S.S. (2009) An Analytical Way to Find Influencers on Social Networks and Validate their Effects in Disseminating Social Games. *Proceedings of International Conference on Advances in Social Network Analysis and Mining* (*ASONAM*), Athens, 20-22 July 2009, 41-46.

https://doi.org/10.1109/ASONAM.2009.59

[68] Li, Q., Zhou, T., Lü, L. and Chen, D. (2014) Identifying Influential Spreaders by Weighted LeaderRank. *Physica A*: *Statistical Mechanics and Its Applications*, **404**, 47-55. https://doi.org/10.1016/j.physa.2014.02.041

[69] Cossu, J.-V., Dugué, N. and Labatut, V. (2015) Detecting Real-World Influence through Twitter. *Proceedings of* 2015 2*nd European Network Intelligence Conference*, Karlskrona, 21-22 September 2015, 83-90. https://doi.org/10.1109/ENIC.2015.20

[70] Sun, Q., Wang, N., Zhou, Y. and Luo, Z. (2016) Identification of Influential Online Social Network Users Based on Multi-Features. *International Journal of Pattern Recognition and Artificial Intelligence*, **30**, Article ID: 1659015. https://doi.org/10.1142/S0218001416590151

[71] Zhuang, K., Shen, H. and Zhang, H. (2017) User Spread Influence Measurement in Microblog. *Multimedia Tools and Applications*, **76**, 3169-3185. https://doi.org/10.1007/s11042-016-3818-z

[72] Tan, C.W., Yu, P.-D., Lai, C.-K., Zhang, W. and Fu, H.-L. (2016) Optimal Detection of Influential Spreaders in Online Social Networks. 2016 *Annual Conference on Information Science and Systems* (*CISS*), Princeton, 16-18 March 2016, 145-150. https://doi.org/10.1109/CISS.2016.7460492

[73] Xia, Y., Ren, X., Peng, Z., Zhang, J. and She, L. (2016) Effectively Identifying the Influential Spreaders in Large-Scale Social Networks. *Multimedia Tools and Applications*, **75**, 8829-8841. https://doi.org/10.1007/s11042-014-2256-z