



A Survey on Social Media Influence Environment and Influencers Identification

Feriel Gammoudi^{1,2} · Mondher Sendi² · Mohamed Nazih Omri²

Received: 10 May 2022 / Revised: 10 September 2022 / Accepted: 17 September 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2022

Abstract

Social media users can be influenced directly by their close relationships, such as their friends, family, and colleagues. They can also be influenced by those who follow them through shared information, goals, news, and opinions. Generally, an influencer is someone who entices an influence to do the same action, make the same decision, or change their behavior. He can also communicate information, ideas, and thoughts to multiple users. There are many tools to identify influencers. It can not be found simply through their big follower number or their shared media number. Thus, influencer identification is one of the essential tasks in social media research. Several approaches and metrics have been proposed in the literature to identify influencers. In this article, we explored the issue of identifying social media influencers while providing a generic view of social media influence. First, we presented a literature synthesis on the influence of social media. Then, we categorized the works and illustrated the leading solutions in literature to identify influencers in social media. A discussion and suggestions for potential future directions in this area accompanied this presentation. We believe these briefings are critical to resolving the issue discussed in this article.

Keywords Social media · Influencer · Influencer identification · Influence · Finding influencer · Users' interest

1 Introduction

1.1 Context

Digital networks have a huge impact on people's daily lives, leading dozens of new researchers to study this rapidly growing issue that extends throughout multiple sub-domains. As social networks become more prominent, a researcher's first logical thought upon analyzing these networks is how to get specific information from the data. That makes marketing applications explore social media to understand user trends and offer new products. Military applications have analyzed

them to detect terrorist groups and their further socialization Cialdini and Goldstein (2004). The health and biomedical fields Sendi and Omri (2015) have also benefited from social media analysis by exploring discussions on health topics and analyzing patient behaviors. Analysis can also predict relationship quality, event detection and characterization as good or bad Iraklis and Williams (2020) anticipating the number of followers they might have Imamori and Tajima (2016), enhancing interaction activity and user similarities Xiang et al. (2010). And by combining temporal information with a social network structure, user interests Bao et al. (2013) are predicted.

As the researchers' results found, the network structure proves to be of utmost importance. A graph with nodes and edges, where each edge symbolizes the connection between two users, whereas each node represents one user, can be used to model a social network. Users have a personal web page and connect with different friends to share content and interact with each other, and build social relationships by following others or adding a theme as a friend.

In associated with social influence Peng et al. (2018) an influencer entices an influence to do the same action or make the same decision, and there can also be no uncertainty Peng

✉ Feriel Gammoudi
GammoudiFeriel@gmail.com

Mondher Sendi
Sendi.Mondher@gmail.com

Mohamed Nazih Omri
MohamedNazih.Omri@eniso.u-sousse.tn

¹ Monastir Faculty of Science, University of Monastir, Monastir, Tunisia

² MARS Research Laboratory LR17ES05, University of Sousse, Sousse, Tunisia

et al. (2017). Users might well be instantly influenced by their true and intimate relationships (friends, family, classmates, coworkers, etc.) while they spend more time with them, and by famous content creators that they follow, by consulting their shares daily: actuality, goals, news, and opinion. Thus, users are sharing influence while sharing media and posts. The information and knowledge management community has recently attached importance to the topic of identifying significant users from social media as “the source of influence”. Although studies have concentrated on measuring levels of ability to identify knowledgeable users and determining levels of influence to identify influential users. The process starts by analyzing the influence of the nodes and then progressively maximizes the influence of an initial set of nodes to obtain the influential nodes. It is a challenging task that has the potential for considerable usefulness in many applications.

Several important works in the literature deal with influence and many new works have recently been attracted by identifying important and significant influencer nodes. In this paper, we will deal with the social media influence environment, influencer identification, and important nodes, and we will classify important existing works.

The following section discusses the main related surveys cited in the literature.

1.2 Related surveys

Many important works dealing with social media analysis have been proposed in the literature. We are going to classify them according to their state of the art and their metrics. The first social media analysis technique Adedoyin-Olowe et al. (2014), Tabassum et al. (2018) was classified as an unsupervised, semi-supervised, or supervised learning method. The SM community detection Azaouzi et al. (2019) and social media structure discovery have two types of structure networks: static or dynamic according to space and time. Some others worked to down the social media users’ relationships as a graph with nodes and edges. Authors Bian et al. (2019) reviewed, analyzed, and classified some main literature works into two main categories: the most influential nodes in the graph and the most significant nodes. Works dealing with the inferring of user interest profiles have been classified in Piao and Breslin (2018). Another target user personality has been detailed by its authors in Andreassen et al. (2016). Depending on the sort of machine learning utilized in each approach, the information transmitted could be rumors Alzanin and M.Azmi (2018). The authors proceeded by dividing the problem of rumor detection into three different categories: hybrid approaches; supervised approaches; and unsupervised approaches. There are also important works dealing with the analysis of social media influence, such as the works detailed in Peng et al. (2018); Sun and Tang (2011); Peng et al. (2017). The authors of the

contribution presented in Peng et al. (2018) classified several notable works using switch algorithms and models. In Sun and Tang (2011), the authors also described various methods and algorithms for calculating measures related to social influence. An important new work Azaouzi et al. (2021) deals with the problem of influence maximization under privacy protection. They divide models into two categories: group node-based models and individual models. And, by proposing different algorithms, different methodologies, and also diverse frameworks, they demonstrated that one key solves the problem of maximizing influence at the individual level Kempe et al. (2003). The perspectives of influencer marketing, their content strategies, and the attractiveness of their sponsored recommendations are reviewed to promote the marketing tactic. The three research factors used by authors to categorize papers were the source, message, and audience Hudders et al. (2021). Social network applications are becoming highly interested in the topic of community detection Plantie and Crampes (2013).

There is much work dealing with the influence of others on social media. We aimed to classify some of them subjectively according to their aims and the taxonomy used in classification. Table 1 below summarizes the related surveys.

To the best of our knowledge and after these related surveys studies, no one proposed a classification of influencer node. Moreover, no work has been found to deal with the global problem of social media influencer identification. By comparison with the main related surveys, such as Bian et al. (2019); Adedoyin-Olowe et al. (2014); Tabassum et al. (2018); Piao and Breslin (2018); Alzanin and M.Azmi (2018); Hudders et al. (2021); Andreassen et al. (2016); Peng et al. (2018); Plantie and Crampes (2013) presented previously, our study presents the following major differences:

- We are not only dealing with influence maximization or significant node identification. Our aim is the identification of many types of social network influencers.
- Unlike some existing surveys, dealing with only one type of social network, that of Twitter, our study exploits different types of social networks, such as micro-blogging, egocentric networks, Google+ dataset, Facebook, Amazon, DBLP, and Wiki.
- Our study is characterized by two important points: exhaustiveness and recent bibliographical references. Indeed, our study represents coverage of different fields of application.

1.3 Motivation and contribution

1.3.1 Motivation

Numerous important works in the literature deal with user interests and influence, which have changed over time and have played a vital role in all analysis processes. Thus, we

Table 1 Classification of the main surveys

Reference	Aims	Taxonomy(ies) studied	Finding
Piao and Breslin (2018)	Inferring User Interests	- Data Collection - Data Representation - Data Construction - Data Enhancement	Both questionnaires and extrinsic evaluation procedures have similar levels
Alzanin and M.Azmi (2018)	Detecting rumors	- Unsupervised approaches - Supervised approaches - Hybrid approaches	There is still a critical need to include many languages
Sun and Tang (2011)	Social influence analysis models and algorithms	Comprehensive survey	An important and challenging research area
Adedoyin-Olowe et al. (2014)	Social Network Analysis Using Data Mining Techniques	- Unsupervised Methods - Semi-supervised methods - Supervised methods	The usefulness of data mining techniques in locating important content and information among the massive amounts of data created
Tabassum et al. (2018)	Social network analysis	Comprehensive survey	Still a challenge for many of the metrics that need to traverse the entire or most of the graph on every update
Andreassen et al. (2016)	The relationship between addictive use	Behavioral addiction	It is necessary to conduct an additional study on these understudied interactions
Plantie and Crampes (2013)	Community Detection	- Graph partition - Hypergraph - Concept graphs or Galois lattices	There should be more approaches developed, and software tools to support them are anticipated to follow
Azaouzi et al. (2019)	Community detection	- Static network - Dynamic network - Real-world datasets - Synthetic datasets	Important topic
Bian et al. (2019)	Identifying Top-k Nodes in Social Networks	- Influential nodes - Significant nodes	- The work in this area is quite limited. - The research area of mining top-k nodes in dynamic networks is still relatively new
Peng et al. (2018)	Influence analysis	Comprehensive survey	Social networks are on the horizon, and social influence analysis is a crucial field to address needs
Azaouzi et al. (2021)	Influence maximization Models	- Individual models - Group node-based models	Many more problems and challenges will show up during the progress of the social network journey
Hudders et al. (2021)	Social media influencers' strategy	Comprehensive survey	- Research on influencer marketing is flourishing - Continuing research on the subject is estimated to grow
Our Survey	Influence and identifying influencers	Comprehensive survey	- Interesting issue - The amount of ongoing research is anticipated to rise

have dealt with social media strategy, the first user influencer, to simulate the real environment of users.

The social media strategy makes users more dependent: content is organized and categorized, and switching algorithms, user activity, and past choices increase the usage rate. One of its strategies is that publication occurrence depends on users' last choices and their degree of importance, the rate of publication consulted, and users' litigation. These networks use cookies and algorithms to collect

information about them and their behavior. Other websites consult, search, and try to identify the users' interests. When the latter is looking for something, it starts appearing as a suggestion on social media homepages. This further influences brands to consume their offer. User interest was also guided by photos, videos, and information inspired by his friends. On the homepage are their trending topics, most viewed videos, and their latest picks; similar

and international trending topics; international action or famous events; and personal activity.

Social media allows users to follow each other and share details about their activities and postings. They can like each other's publications, photos, videos, opinions, and disputes. It allows the user to see similar consulted publications, and it offers the possibility to chat and create direct communication and influence. Their friend or an influencer can be inspired by seeing the publication. That is why we thought that social media users were sharing influence.

All of those influence the user's choice, use, ideas, focus, and personality. Users prefer personalized and guided services that are matched by their choice of searches and hope. So we are dealing with the influence and influencers on social media networks.

1.3.2 Contributions

The main contributions of this article can be summarized in the following points:

- First, we dealt with the social media influence environment to study the general concept of our aims.
- Secondly, our principal contribution was influencer identification and important nodes, followed by classifying important works dealing with them and reviewing them.
- Finally, in a third step, new research directions concerning our survey review are discussed.

1.4 Paper Organization

The rest of this work is organized as follows: In Section 2, we presented the review methodology adopted in this work by providing the information sources and the research questions. In section 3, the social media influence background is presented and discussed. Section 4 is devoted to detailing an overview of social media influence. Then, section 5 contained a literature synthesis on the social media influence that we have treated in this work. Section 6 presented the discussions and research challenges that are synthesized at the end of this research work. We concluded the study by proposing several research challenges in Section 7.

2 Review methodology, sources of information and research questions

2.1 Review methodology

We have assumed a systematic methodology to elaborate on this current work on social media influence. The following is a description of the review approach used, inspired by Kitchenham and Charters (2007); Helali and Omri (2021);

Souiden et al. (2022). The review procedure is a series of processes that make up the review protocol. We begin by providing the various information sources and then the search and selection criteria for the main bibliographical references that we used to undertake this study. Next, we present the questions section that we have to provide answers to throughout this literary work.

According to our objective, we do not aim to create a proper list dealing with social media influence but to introduce some important work. We focused on works from reputable publications. The initial research was based on keywords from recent works. After that, we filtered our search results by choosing according to a few norms, such as the publication date and/or writer competence. If the same, we find a similar approach or idea. Then, based on a full-text screening, title, and abstract, we obtain our primary studies list. Finally, to ensure the quality of the assessment form, we choose several articles arbitrarily based on the previous stage's selected results.

2.2 Sources of information considered

Concerning the bibliography used, we proceed to search various information sources to lead to the realization of this study. To choose sources related to our focus problem, we selected research articles from journals, conference proceedings, books, and magazines. Thus, we considered the following databases in our research: IEEE Xplore Springer¹, ScienceDirect², Scopus³, ACM Digital Library⁴, Taylor⁵, Francis⁶, Google Scholar⁷.

We also screened the majority of related high-profile conferences, such as SIGKDD, ICDT, SIGMOD, ICML, WWW, VLDB, ICDE, and EDBT, to find out about recent work. Figure 1 shows that papers reviewed percentages from different types of resources.

2.3 Research questions

In our research, we are going to answer some proposed questions that are identified in this current section. Therefore, this part focuses on question determination, as below. Then, we will answer those questions through the research review.

- RQ1 : Who can influence social media users?

¹ <https://link.springer.com>

² <http://www.sciencedirect.com>

³ <https://www.scopus.com>

⁴ <https://www.acm.org/digital-library>

⁵ <https://www.taylorandfrancis.com>

⁶ <http://ieeexplore.ieee.org>

⁷ <https://scholar.google.co.in>

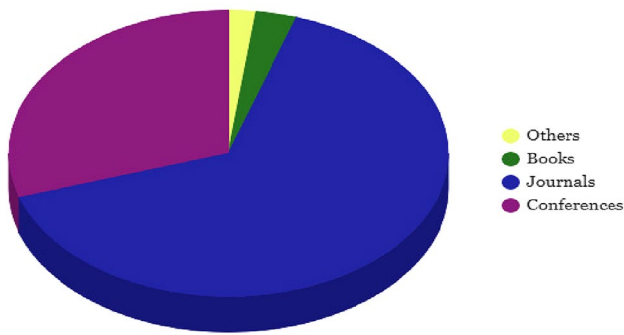


Fig. 1 Articles from different types of sources

- RQ2 : Which is the way to detect influencers?
- RQ3 : How do they influence social media users?

3 Social Media Influence : background

3.1 Social Networks representation

Social media networks can be represented as a set of profile sites Aggarwal and Subbian (2014) or virtual communities. Their structure, by default, is considered static, with a static topology and information. After that, it can be transformed into dynamic networks at a specific time t if new nodes and edges are added. Networks have a specific topology Peng et al. (2018): internet sites where users connect and communicate with others; and mobile social networks. According to the contribution Bian et al. (2019), the authors show that we can model a social network by a graph $G = \{V, E\}$ where $V = \{v_1, v_2, \dots, v_i\}$ represents a set of nodes and $E = \{e_1, e_2, \dots, e_j\}$ represents a set of edges. In the network graph, actors are denoted by V , and social interactions and relations between them are denoted by E . Edge measures relay the influence built on the concepts of a pair of nodes. Each node v_i can be active or inactive, and the diffusion of influence can propagate from the active node to its neighbor, and the links correspond to social relations.

3.2 Influence on Social Media

Social influence is a concept that includes a wide range of phenomena, such as:

- Socialization is the process through which individuals acquire the knowledge, skills, attitudes, values, norms, and appropriate actions of their community Shahr et al. (2019), which powerfully directs behavior in private settings as well Cialdini and Goldstein (2004).
- Obedience and authority Individuals are frequently rewarded for behaving under the opinions, advice, and

directives of authority figures Cialdini and Goldstein (2004).

- Compliance professionals are forever attempting to establish that they are working for the same goals as they are, in essence, their teammates Cialdini (2001).
- Conformity refers to the act of changing one's behavior to match the responses of others Cialdini and Goldstein (2004).
- Persuasion, the ability to use its effect, is frequently a crucial element of success Cialdini and NJ. (2002).

Social networking users tend to act like their “friends” or “neighbors” Singlaand and Richardson (2008); Sun and Tang (2011). This phenomenon can be summarized in the following points:

- Social Influence: Social media users are more likely to imitate their friends' actions. Thus, social influence causes people to endorse the behaviors of others.
- Selection: People who are characterized as close to them are more likely to bond with each other.
- Variables that may be confusing: Other unidentified variables exist, and people may act similarly to each other.

Social influence Cercel and Trausan-Matu (2014) can be defined as the influence of a person or a group of people on each other. This influence is manifested through the imposition of a set of attitudes and behaviors. Considering two users, U_1 and U_2 , from the social network, U_1 has a direct influence on U_2 or indirectly influences the point of view of U_2 , and transmits influence to U_2 . According to Peng et al. (2018), social influence is defined as a level of uncertainty or a binary relationship between one (the influencer) and the other (the influenced) to do the same action or choice. The level of social influence can be quantified by a continuous real number or by a degree of uncertainty (e.g., weaker, weak, strong, stronger, etc.) and can have different values Peng et al. (2017). The fact that U_1 influences U_2 , does not imply that U_2 necessarily affects U_1 .

3.3 Two-step and multi-step communication flow theory

Since relationships are organized around a network of peers, with no centralized entities of control in social media, users are thrown into a global stream of communication Elanor (2013). Social media networks have reshaped communication at different levels, such as mass communication's two-step flow of communication theory Oren (2019) and multi-step communication. The two-step flow communication theory can be used to describe how influence works on social media. According to Ognyanova (2017) in the two-step model, the theory's concept of

personalized content through opinion leaders Oren (2019) and direct media effects are hampered by social interactions and audience selectivity in exposure, perception, and retention. The two-step flow theory was expanded upon by the multi-step paradigm Ognyanova (2017). The multi-step flow paradigm describes the way media and interpersonal influence shape public opinion. Understanding whether someone's opinion was influenced by their social links or if they simply chose social ties that shared that opinion, to begin with, is important in the context of the multi-step flow of communication Ognyanova (2017).

3.4 Influencers identification

Nowadays, social media homepages are limited to showing their users priority on the activities of their friends who are used to communicating with or reacting to their statuses, pics, media, or stories. This task can be studied by utilizing the rich text and user interactions on social media Zheng et al. (2020). Some research deals with the person who influences others, each defines a type and researches its impact on others. Nowadays, social media homepages are limited to showing their users the activities of their friends, who are used to communicating with or reacting to their statuses, pics, media, or stories. Some research deals with the person who influences others, each defines one influencer type and researches its impact on others. In the book, Shah et al. (2018) and according to its authors, influential people play an important role in information sharing. Being the most popular or the first to affirm a new concept are two examples of being influential. They also provide a strategy for identifying network infrastructure pioneers based on a specific topic of interest. People who adopt and introduce new ideas before they are known, influencing others, are known as trendsetters. They are not always famous or popular. In the same way, not all the latest inventors are trendsetters since only a few of them can spread their ideas in their social circles. They also suggest a reliable way to describe the spread of creativity by presenting each issue as a collection of trends that can be applied to many different scenarios. Results from algorithm simulations on a large Twitter dataset show an ability to quantify indirect and direct influence, and early adoption is also an important trait for influencing others.

All this is to distinguish between the trendsetters and the others, who, although having a high degree, only accept trends when they have become well known. Other specified innovation hubs score a lower degree of acceptance of a new idea than follower hubs. Followers Innovation Hubs are trendsetters, while hubs are influencers. In the dissemination of information, influential people have a crucial effect. Having more Wang et al. (2020) followers than others does not mean more interaction. There are several ways to be

influential, such as being the most famous or being the first to implement a new idea. The most influential users Erlandsson et al. (2016); Zareie et al. (2019); Zhao et al. (2019); Sun and Ng (2012) played an important role in spreading the information. Others aim to find influencers Rodríguez-Vidal et al. (2019); Subbian et al. (2014); Harrigan et al. (2021); Tsugawa and Kimura (2018); Kaple et al. (2017); Chia et al. (2021); Harrigan et al. (2021) in different ways. The authors of the works presented in Cervellini et al. (2016); Saez-Trumper et al. (2012) deal with a trendsetter who is not automatically popular or famous but whose thoughts have been broadcast on the network. As a result, they serve the same purpose with a different name Zhou et al. (2019): influence nodes. Identifying influential users has important implications in e-commerce and media Jain and Sinha (2020); Sun et al. (2016); Yang et al. (2019); Sheikahmadi et al. (2017). Another example is influential actors Qasem et al. (2015, 2017), who have followers Probst (2013), are famous on Instagram Jin et al. (2019), and have popular content Ding et al. (2015). This phenomenon causes a significant increase in the size of the social network. In addition, news influencers Alp and Ögüdücü (2018) have been identified as specialists on a specific topic. Table 2 summarizes this work and gives further details.

Trendsetters are people who support and propagate innovative thoughts and have an influence on others before they become famous. Cervellini et al. (2016) propose a methodology for identifying themes in network infrastructure based on a certain topic of interest. And Saez-Trumper et al. (2012) used different ranking algorithms to identify themes. Popular content is by and about famous people on social media. A random walk model was presented by Ding Ding et al. (2015) to quantify the influence of users and the popularity of tweets. Another influencer type is called the prophet, who are knowledgeable bloggers and have a strong capacity to predict the future. Zhang et al. (2015) propose a method for identifying influential blogs that can predict trending hashtags. It is based on keywords or phrases that describe topics or events that are well-known in the community. Influencers are people who have the power to influence someone or something's personality, evolution, or attitude. Authors like Rodríguez-Vidal et al. (2019) used a topic modeling approach to manage the textual signal in tweets after experimenting with numerous signals and machine learning algorithms. Harrigan et al. (2021) are using freely available data from social media APIs, and influential mavens on social media may be discovered. Chia et al. (2021) integrated the social capital and social exchange theories, as well as the social learning theory, to identify ideal Kaple et al. (2017) proposed a method to maximize public participation and build smart cities through the use of social networks. To detect influencers, Pudjajana et al. (2018) employ SNA metrics (DC, CC, BC) and weight on the SNA measurement. A method for identifying network

Table 2 Classification of identifying influencers

Nodes type	Qualification	Identification Process	Related research	Data	Results
Trendsetters	Support and propagate influenced thoughts before becoming famous	- Topic of interest. - Ranking algorithms	- Cervellini et al. (2016) - Saez-Trumper et al. (2012)	- Yelp Dataset - Twitter	- Performs better - The ability to locate a large fraction of trendsetters
Popular Contents	Contents of Popular and Famous People	Random walk model	- Ding et al. (2015)	Twitter	Better than PageRank method
Influencers	Influence personality or attitude	- Topic modeling approach manage textual signal and machine learning algorithms - Using freely available data - Integration of the social capital and social exchange theories - A method to maximize public participation and build smart cities - SNA metrics (DC, CC, BC) - A method based on their social capital value. - Random sampling - Greedy algorithms	- Rodríguez-Vidal et al. (2019) - Harrigan et al. (2021) - Chia et al. (2021) - Kaple et al. (2017) - Pudjajana et al. (2018) - Subbian et al. (2014) - Tsugawa and Kimura (2018) - Sunil and Lingam (2019)	- Twitter - Twitter - General - General - Hoax dataset. - DBLP network - Twitter-follow - SNAP	- Perform above the average - Decision-makers - The structural and relational dimensions influenced SMIs' propensity. - Good performance - Proved - Outperforms PageRank, PMIA and Weighted Degree baselines up to 8 % in terms of precision, recall, and F1-measure. - Possible benefit - Optimal - Outperform
Prophets	Knowledgeable and strong capacity to predict the future	Keywords or phrases describing topics or events	- Zhang et al. (2015)	General	Optimal late-time
Influential nodes	Increase an interest propagation process's asymptotic reach.	Centrality metrics	- Zhou et al. (2019)	General	-The most followers or the highest number of tweets - Proved - Outperforms (speed and capacity) - More powerful
Influential users	Authoritative actors	- WCI integrates 10 different elements into two - Top 1 % following growth. - Neighbors' adoption delays, and the spreading - A multi-features model-based and ranks impact using the Page-Rank concept	- Jain and Sinha (2020) - Yang et al. (2019) - Sheikahmadi et al. (2017) - Sun et al. (2016)	-Twitter - Twitter - Twitter - Sina Weibo	
Most/High Influential Users	The top-k users	- Degree Centrality and Page Rank Centrality - An algorithm is interested in the messages - Interactions - High influence user discovery algorithm	- Erlandsson et al. (2016) - Zareie et al. (2019) - Sun and Ng (2012) - Zhao et al. (2019)	- Facebook - General - Facebook and Twitter - Microblogging	- A lower execution time - Effectiveness - Proved - Best performance
Micro-influencers	A special kind of influencers, who are harder to find, less famous but with a higher engagement power over their communities	Automatically approach and highlight personality traits and community values, by analyzing their writings	Leonardi and Monti (2020)	Twitter dataset	Best performing - Impact

Table 2 (continued)

Nodes type	Qualification	Identification Process	Related research	Data	Results
Influential Actors	Tweets generate an enormous number of retweets	- Novel influence degree. - The attractiveness model is defined with the T measure	- Qasem et al. (2015) - Qasem et al. (2017)	- Asterisk -Twitter	-Proved -Optimal
Topical influencers	Experts on a given topic	Related features and network feature information. -Based on the language attention network and influence convolution network -Unified hypergraph to model users, images, and various types of relations	- Alp and Ögüdüci (2018) - Zheng et al. (2020) - Fang et al. (2014)	Turkish tweets -Twitter -Twitter	- More efficient. - Comprehensive -Effectiveness and improve the performance
Airline Influencers	influencers relevant to the brands of several airline companies	SNA-based approach	Izdihardian and Ruldeviyani (2021)	Tweets of Indonesian Airline	Important
Instafamous	Instagram celebrities	Test the effects of celebrities	- Jin and Ryu (2020)	Instagram	Discusses theoretical implications
Lead or Follow Consumer Profiles	Mobilizers and propagators Consumer of products or services	Systematic literature review Analyzing noisy social media data and messages	- Florian, 2013 Probst (2013) Hernandez et al. (2013)	OSN data set Twitter	Discussion A 10% random sample
Partisan Slant	Political people	A natural language processing algorithm to analyze at scale the linguistic markers	Karamshuk et al. (2016)	Twitter	Inferred with an accuracy of 60-77%
Pathogenic	As terrorist supporters exploit large communities of supporters for conducting attacks on social media.	A classification algorithm to classify accounts	Alvari et al. (2018)	Twitter	F1-score of 0.6 and via supervised learning identified 71% of the PSMs
Central nodes	Centrality is the most prominent measure that shows the node importance from the information flow standpoint	DICeNod, the compressive sensing-based framework	Mahyar et al. (2018)	Real networks from SNAP	Very well in terms of number of correctly identified central nodes
Influential mavens	Influencer marketing	user network, user behaviour, message readability, and message structure	Harrigan et al. (2021)	Twitter	Effectiveness and efficiency
The Most Influential Spreaders	spread of important information (highly beneficial or alarming if false content)	UACD, a novel method combining both user-specific and topological information	Adnan et al. (2022)	Amazon EC2	Scalable and can process a very large network
Critical nodes	Whose removal reduced the flow of hate	Analysed the graph characteristics then classified as containing evidence of hateful content	Alorainy et al. (2022)	A collected and large dataset	Effectiveness

influencers based on their social capital value was proposed by Subbian et al. (2014). Switch Sunil and Lingam (2019) most of the currently used algorithms created to address these issues have the greedy algorithm as their foundation. And the two-level approach was proposed (SI) for an epidemic model for maximizing the influence spread and a multithreading approach for implementation. Tsugawa and Kimura (2018) used degree, betweenness, proximity, PageRank, and the k-core index, and used the common sample methodologies SEC, BFS, DFS, and random sampling to find influencers in different social networks. Switch topic Alp and Ögüdücü (2018) defined the topical influencers as experts on a given topic. And to identify the theme, an approach that combines related features and network feature information was used. It is widely acknowledged that the distribution of influence varies over topics. The work of Zheng et al. (2020) is based on the language attention network and influence convolution network to detect on-demand topic-specific influencers. Another work dealing with Topic-Sensitive Influencer Fang et al. (2014). They used a unified hypergraph to model users, images, and various types of relations. Influencers of airline services Izdihardian and Ruldeviyani (2021) are also detected on Twitter and the results showed their higher influence. Influential nodes in complex networks are often those that increase an interest propagation process's asymptotic reach. Zhou et al. (2019) a thorough evaluation of centrality metrics based on their capacity to detect early and late influencer nodes. Lead or follow the influence of mobilizers and propagators in networks and society funded by Probst (2013) using a systematic literature review to find relevant publications. Influential users are authoritative actors and are a term used to describe influential users. The work of Jain and Sinha (2020) used WCI (Weighted Correlated Influence), a new concept that integrates the relative influence of 10 different elements into two different feature sets: timeline and trend specific. Yang et al. (2019) developed the component study of the characteristics of Houston's top 1% of Twitter users in terms of the following growth. Another important work by Sheikahmadi et al. (2017) where weights are allocated to neighbors based on their adoption delays and the spreading influence of nodes is identified. Via a multi-features model, Sun et al. (2016) and based on a user identification solution with a lot of clouts. The model investigates if a user's attributes, such as follower count, can be used to infer whether he or she is an influential user, and then ranks influential individuals based on their impact using the Page-Rank concept. Others aim to find the most influential users, or "top k" users, who are considered the most influential. Erlandsson et al. (2016) found the Degree Centrality and Page Rank Centrality and compared the theme to the outcomes of association rule learning. Zareie et al. (2019) proposed the idea of a user behavior analysis-based high-influence user discovery algorithm (HIUD). Also, influential actors are individuals whose

tweets generate an enormous number of retweets, resulting in a larger social network, proposed in 2015 by Qasem et al. (2015) find via a method for detecting influential people utilizing a novel influence degree. Then, Qasem et al. (2017), an extended method for identifying important actors, was formed based on the attractiveness model defined with the T measure.

Other works have defined special influencers that can be an inspiration for many important other works by trying to identify and find those influencer nodes on social media networks. MARWICK (2013) wrote a book that describes content creators who enjoy fame and influence others by using social media to maintain an active communication network. Other works could use the larger social graph or friends list and broadcast audience recordings to Marwick, who proved its importance. In the work Enke and Borchers (2019), the authors described three types of influencers: digital opinion leaders, those who are likely to have an impact on others in their immediate area, and content creators, who are people who develop content for social media using their talents to create social media content. They also defined moderators, who initiate interaction procedures and utilize their position to try to lead and continue conversations, and they proved the importance and strategy of influencer communication, in which social media influencers are performed or addressed within tasks that are strategically important to the objectives of business goals. Campbell and Grimm (2019); Campbell and Farrell (2020) aimed at users with large communities of followers. Like celebrity influencers and mega-influencers, last but not least, are individuals who have already earned 1 million or more followers on social media, have seen considerable follower growth, and have gained celebrity status through a well-established skill. Recording Campbell and Grimm (2019) authors' social media influencers' posts on their social media channels in exchange for compensation. In their next work, they affirmed that influencers can be simultaneously social media personae, celebrities, experts in particular fields, and peer consumers. They add another definition of influencers and micro-influencers, whose ads are indistinguishable from other consumer-generated content. And also, they defined influencers as individuals who post to their social media accounts in exchange for compensation and as consumers paid to endorse products and promote them to their followers. They affirmed that micro-influencers, Campbell and Grimm (2019); Campbell and Farrell (2020); Leonardi and Monti (2020) or influencers with relatively small online followings, are more effective, often implying enjoyment or simply usage than other influencers and celebrities.

Using social media to identify influential users is a complex process, but one that is extremely important for several Zhang et al. (2015) applications. Previous research has determined skill levels for detecting knowledgeable users and

influence degrees for recognizing influential users. Generally, they mainly depend on the analysis of textual content but also on links to the structure of social media networks. Bloggers use the blogosphere as a platform to post articles, communicate ideas and discuss views. Blogosphere data is dynamic and changes with information updates over time. In the work presented in Zhang et al. (2015), the authors attempted to find bloggers who anticipate future popular trends, known as prophetic bloggers. In this work, the authors examined the chronological characteristics and content of blog data and proposed a method to analyze the ability of bloggers to predict buzzwords. Bloggers are judged on the types of words used in their blogs. For example, four recent, related, frequent, and/or thorough factors are considered for this goal: content similarity, post-precocity, buzzword coverage, and frequency of entry. In the work of Zhang et al. (2015), the authors stated that by inferring buzzwords, they detected buzzwords, which allows the assessment of bloggers' buzzword prediction ability. The contribution proposed by Bian et al. (2014a), was interested in the detection of epidemics. To discover the spread of a virus as soon as possible, they have selected many social network nodes that they consider most relevant. To maximize the benefits of companies' products and services, consumer profiles are detected and analyzed Hernandez et al. (2013). By analyzing a dataset from Twitter, the linguistic indicators of partisan conflict in mainstream and social media during political upheaval are explored Karamshuk et al. (2016). Others deal with pathogenic social media accounts, such as terrorist supporters exploiting large communities Alvari et al. (2018) and critical nodes, Alorainy et al. (2022) and a high detection of top-k central nodes Mahyar et al. (2018). Identifying the users who are most influential in spreading information can aid in the creation of effective tactics Adnan et al. (2022).

4 Overview of Social Media Influence

4.1 Users' interest

The degree of influence a social media user has on their friends, as well as their desire to spread additional information, can be detected from interest similarity. Authors have submitted that people users are more affected and influenced by their friends' opinions than by other influencers Domingos and Richardson (2001); Staab et al. (2005). In Xie et al. (2014), the measure of similarity between two users is based on a random walk distance based on tags and image correlation and was examined in folksonomy data. It outperforms proof in terms of user profile enrichment. Generally, social media users demonstrate Bian et al. (2014b) different interests and hold different expertise levels for numerous kinds of interests,

such as sports, music, history, and so on. Authors Sendi et al. (2017) assumed that users' interests were discovered from uncertain information. Another way to categorize users' profiles Mabrouk et al. (2020, 2018) as an outcome, the culture of the user's interest profile will be critical for a full analysis of their influence on their friends, diffusion actions, and future action prediction.

4.2 User's relationship

The main reason that social networks influence users' behavior changes is their relationship. As users' relationship strengths are not equal, social media users do not have the same degree of friendship and intimacy. Relationship strength can be detected by interaction comments, likes, chats, etc. Previously, social network studies centered on binary friendship relationships. Switch authors Xiang et al. (2010) can lead to a wide variety of relationship strengths (for example, mixing acquaintances and best friends) to estimate relationship strength. They create an unsupervised model based on user similarities and interaction activity, weak and strong relationships. Their model represents the full spectrum and infers a continuous-valued relationship strength for links. For the inference, they used a coordinate ascent optimization approach to create their link-based latent variable model. They have proved and evaluated it on Facebook and LinkedIn. The authors Chader et al. (2017) did not accept the binary relation. They had related weights that recorded the strength of people's relationships; they tested the effectiveness of the CoBSP process. The motive comes from the belief that those with the closest relationships with users can give more information about them. In the work of Zarrinkalam et al. (2017, 2019), they detect user interests by focusing on social relationships, user-generated content, and temporal factors.

4.3 Interests change

After taking into account the users' utterances, gender Ouni et al. (2022b), their activeness, and vulnerability to their friends' influence, the work of Budak et al. (2014) proved that interests vary during life. They established an inference algorithm that strikes a balance between the latest estimates of interests and the old estimate of interests, and using Twitter data, they funded a precision of 0.9 for the top-5 interests. Users have a variety of interests, such as music, history, sports, and so on Bian et al. (2014a). The degree of influence he has on his friends is determined by his interest in distributing new information and must be varied for different categories. The interest profile learning is decisive for a complete analysis of a user's impact on his contacts, spread actions, and future action prediction. Another idea Wang et al. (2018) is to use the UNITE systemic framework with interesting

extraction, which considers both text and information. Their approach to experimental results on Sina Weibo. An important work Sendi et al. (2019) aims to derive and monitor on Twitter the temporal interest topics and also handle the user's interest, social various dimensions, and dynamical characteristics. Moreover, over time on social media, through user activities' frequent tags, users, and organizations from their social media posts Shah et al. (2018), the evolution of one's attitude through time is observed. The suggested approach also incorporates the use of the semantic web, particularly the DBpedia ontology, which analyzes the content of tweets in the user's Twitter feed.

4.4 Infer interests

To deduce users' interests, authors have taken into account their communication, their activity state, and their sensitivity according to the influence of their friends. A probabilistic model of social data has been proposed by Budak et al. (2014). This model is unsupervised and well-scalable with many users and interests. The authors of Han et al. (2015), were interested in determining the similarities of interest given to two users, without having any idea about their interests. They showed that similarities between people depend on similar demographic information and having more common friends. Attribute inference can be performed by mixing social graphs and Jia et al. (2017) behaviors, by first collecting public data and then using machine learning techniques to estimate the private attributes of certain target users. They analyzed user behavior to determine the probability that all users have a given property. Many applications, such as targeted advertising and news suggestions, rely on user locations, which can help solve the problem of profiling users' Ouni et al. (2022c) home locations Li et al. (2012). The authors propose a unified discriminatory influence model UDI and overcome the challenge of rare signals in a unified probabilistic framework, combining user-centric and social media data. They produce local and global locations with their prediction methods. Important work has also been proposed in Utz (2010), their authors found that the target's acknowledged popularity, community orientation, target's self-generated data, the affectivity of the target's contacts, and the volume of friends all have an impact on social attractiveness. So users can create accounts with photos and detailed descriptions to use as a starting point for building an impression. In addition to self-provided data, social media offers two other types of data: the target's friend data and system data, such as the number of friends. Another interesting work has been proposed in Xie et al. (2014) where the authors aimed to improve various user tags for a simple image by establishing a collaborative description and making them more meaningful and useful to users.

4.5 Emerging and trending topic

The work proposed in Bao et al. (2013) was occupied by trending topics. It could be related to important news or subjects being discussed by a large number of microbloggers. The authors supposed that publishing posts on a popular topic proved the user's interest. The latter are replicated in posts published on the same trending topics. In the work presented in Dang et al. (2016), the authors aim at emerging topics and are concerned with the substance of important emerging events such as regulatory enforcement, election campaigns, natural disasters, and traffic accidents. Dynamic Bayesian networks are used to uncover developing topics within their methodology. For this, they started the topic diffusion characteristics in the early phase, as well as non-emerging topics with several topology features, and they compared emerging and then, in a given time interval, generated an emerging keyword candidate term list based on term frequency. Next, create a DBN-based model for each candidate to calculate the probability of the candidate emerging as a keyword based on the joint conditional probabilities of the specified features.

5 Social Media influence and literature synthesis

In this section, we proceed to an illustration and a classification of the main relevant works in the literature that are related to the problem treated in this work. The Table 3 presents a summary of the works covered and discussed in this survey. We can see that the majority of previous studies have dealt with the goals and interests of users. Many other studies have addressed the influencers, reasons for interest, and change of all new users.

This Table 3 also presents the approaches used, the experimental results, as well as the years of publication, the authors, and the methodologies adopted. The different approaches to this survey are listed with their concepts. The table also shows the average scores and proof.

For example, in 2013, Bao et al. (2013) propose a PMF-based method to predict user interests by combining temporal information and social network structure. In this work, the authors confirmed the accuracy of user interest predictions. Then, in 2014, Bian et al. (2014a) focused on the prediction of trending messages and diffusion participants. Three types of influence have been determined: epidemic influence, interest influence, and social influence, which has proven its superiority. In 2015, Han et al. (2015) were interested in inferring similarities in interests using user demographic information, friendships, and interests. The proposed model was tested on a Facebook dataset. The authors were able to conclude that people with similar demographic

information (e.g., age, location) or more mutual friends tend to have more similar tastes. Another interesting work is that of Xiang et al. (2010). This work consists of modeling the strength or weight of the relationship using a latent variable model on Facebook and LinkedIn. The simulation results of this model showed better autocorrelation and classification performance compared to the results obtained using different raw data elements. The contribution proposed by Zarrinkalam et al. (2017, 2019) leveraged influencers' interests in active social media users. By using a graphical description model with three types of information, this model showed its performance. Important work was also proposed by Shah et al. (2018), where the authors focused on modeling user interests through the social media network graph and the semantic web. This model has also proven its effectiveness. In the same year, i.e., 2018, Wang et al. (2018) used both content and network structure to extract interest from microblogging and significantly outperform basic methods. Another unsupervised approach that allows inferring user interests from microblogs has been proposed by Budak et al. (2014). The most direct measure of their inference technique is quality. For the top five interests, this model yielded an accuracy of 0.9. Zhang et al. (2015) targeted the chronological and content aspects underlying blog data, and a method of analyzing the power of bloggers to predict hashtags has been implemented and tested. This technique has also shown its performance. In 2017, Chader et al. (2017) show that targeted friendships are not created equal. A methodology for detecting emerging trends has been proposed by Dang Dang et al. (2016). Their DBN model is based on DBN and uses an early detection strategy based on a dynamic Bayesian network for emerging topics in microblogging networks. They choose features among the topological attributes of the subject distribution and use dynamic changes in the conditional probabilities computed by DBNs to recognize new trends. Jia et al. (2017) proposed an AttriInfer model to infer user attributes in online social networks using Markov random fields. They incorporated behaviors and social graphics, benefited from both positive people and people with negative training experience, and scaled across large online communities. They proved that the optimized version of AttriInfer is much more flexible than the fundamental version. Li et al. (2012) proposed their model called "Home Locations Using a Universal and Discriminative Influence Model," which proved its effectiveness. Saez-Trumper et al. (2012) tried to find trendsetters in news networks by proposing a robust method. They showed their ability to locate a large portion of trendsetters, regardless of the top 10% of trend adopters, in authentic situations. In their contribution, when nodes with high degree trends are late, the temporal corrosion function diminishes their impact on outcome ranking and helps to underlie variations in the behavior of TS and PR outcomes over time.

Rather than using the tags or just the photos, they used both user-generated tags and image correlation. The Multifaceted Folksonomy Graph (MFG) was proposed by Xie et al. (2014). The Multifaceted Folksonomy Graph (MFG) was proposed. Recently, Wang et al. (2020), examined the impact of influential leaders on the spread of famous games: Sina Weibo's Travel Frog and assessed the spread trends, as well as the growth of KOL group networks and keywords UGC. They show that the information propagation event continues via periods of robust and slow forward growth, reaching its peak. In the same year, Mabrouk et al. (2020) proposed two approaches: the first is a hybridization of ontology and linear SVM, while the second is a hybridization of ontology and FSVM, where they proved that the SVM fuzzy semantic classifier works exceptionally well Mabrouk et al. (2018). A unique weighted correlated influence metric for Covid-19 is used to identify influential users on Twitter and was also proposed Jain and Sinha (2020). In this work, the authors verified that the individual with the most followers or the greatest number of tweets is not always the most impactful. Trend-specific influence metrics are insufficient to identify influential users. Harrigan et al. (2021) attempted to identify influencers on social networks using data widely accessible from the APIs of these networks, and influential mavens have been identified. They found that big social data can be used by decision-makers to identify influential customers. Also, Chia et al. (2021) proposed searching for the best influencer by combining theories of social capital and socialization with social theories of learning. A complete framework for identifying ideal SMIs has been developed. They proved that the effects of trust and social identification on PMIs' desire to share unboxing journals in the cognitive dimension of the social capital theory were found to be insignificant. The cognitive factor did not influence the propensity of PMIs to post unboxing notices in the community, unlike the structural and relational dimensions. Finding the most influential entities within the network has proved to be NP-hard. The proposed approach More and Lingam (2019) provides a balance between influence spread and execution time. Others Yang et al. (2014) deal with uncovering social network sybils in the wild and terrorist community evolution detection Chaabani and Akaichi (2022) proposing an artificial bee colony optimization, and then applying the BCTTC to track terrorist evolution. In the work of , Hodas et al. (2016), an experiment exploring social media usage during disaster scenarios, combining electroencephalograms, personality surveys, and prompts, is proposed. To detect spammer Aswani et al. (2018) using bio-inspired computing and compromised accounts, the authors Bohacik et al. (2017); Ouni et al. (2022a) used an anomaly model trained on the previous login data of users. Another important work Arora et al. (2019) proposes a mechanism for measuring the influencer index across popular social media platforms

Table 3 Literature synthesis of Social Media influence

Reference	Problem	Concept	Network	Proved
Bao et al. (2013)	Predict users' interests	Predict users' interests: theoretical ideas and integration and time information	Micro-blogging	The accuracy of users' interest predictions
Bian et al. (2014a)	Predicting Trending Messages and Diffusion Participants	Epidemic-oriented, interest-oriented, and social-oriented influence	Micro-blogging	The superiority of this method
Budak et al. (2014)	Inferring User Interests	Probabilistic model for people's preferences throughout time	Twitter data	The most direct measurement and the top five interesting have a precision of 0.9
Chader et al. (2017)	All friends are not equal	Closest relationships to the profiled user discover more useful information about him	Egocentric networks	Relevance of their prior hypothesis
Dang et al. (2016)	Emerging topics	Use dynamic changes to recognize new trends.	Micro-blogging networks	Effectiveness
Jia et al. (2017)	Inferring User Attributes	Integrates behaviors and social graphs, positive and negative	Google+ dataset	- AttriInfer outperforms: efficiency and flexibility - AttriInfer's optimized version is more flexible
Shah et al. (2018)	User Interest	Activities are a part of both the user's and the network's actions	General	Quite effective
Han et al. (2015)	Alike People, Alike Interests	Infer users' interest similarity: demographic information, friendships, and interests	Facebook data	More similar likes with similar demographic information (age, location)
Wang et al. (2018)	Oriented interest extraction	UNITE for collecting interests using textual and structural information	Microblog	Outperforms the baseline methods significantly
Saez-Trumper et al. (2012)	Finding Trendsetters	A robust method for rating trendsetters of innovation diffusion	Twitter	- The ability to locate a large fraction of trend-setters. - Nodes having a high degree, the time corrosion function decreases
Xie et al. (2014)	Community-aware	- User-generated tags and image correlation. - The multi-faceted folksonomy graph (MFG)	NUS-wide dataset	Outperforms the preceding ones.
Li et al. (2012)	User Profiling	UDI determines how likely a user can follow others	Big volume of data	Effectiveness
Xiang et al. (2010)	Relationship Strength	Relationship strength assessment using a latent variable model	Facebook and LinkedIn	The graph auto-correlation and classification performance
Zhang et al. (2015)	Inferring User Attributes	Predict hashtags	General	Outperform
Zarrinkalam et al. (2017)	User interest	Inferred interests as a link prediction problem using a graph description model	General	Significantly improved
Zarrinkalam et al. (2019)	User interest prediction	- User interests vary with time. - Prediction	Twitter	Better performance

Table 3 (continued)

Reference	Problem	Concept	Network	Proved
Wang et al. (2020)	The power of opinion	- Propagation trends - Growth of KOL group networks and UGC keywords	Sina Weibo	- Information propagation - Slow growth is hitting
Harrigan et al. (2021) Jain and Sinha (2020)	Identifying influencers Influence measure	Influential mavens Weighted Correlated Influence (WCI)	Twitter Twitter	Decision-makers -The most followers or the highest number of tweets -Trend-specific influence measurements are insufficient
Mabrouk et al. (2020)	Profile Classification	- Hybridization of ontology and linear SVM - Hybridization of ontology and FSVML	General	Semantic fuzzy SVM classifiers perform well
Chia et al. (2021)	Ideal social media influencers	- Combining the social capital and socialization theories with the social theories of learning - Framework identifying ideal SMIs	General	The structural and relational dimensions influenced SMIs' propensity
Yang et al. (2014)	Sybil in the Wild	- Use ground-truth data about the behavior of Sybils in the wild to create a measurement-based, real-time Sybil detector - Characterization of Sybil graph topology on a major OSN	OSNs	-Still act with no explicit social ties -Effective
More and Lingam (2019)	Optimizing time for influence diffusion	A novel methodology based on gradient approach	SNAP and SLNDC	The greedy algorithm only finds local minimum influence spread -Poorly and limited
Arora et al. (2019)	Measuring the Social Media Influencer Index	A mechanism for measuring the influencer index across popular social media platforms	Facebook, Twitter, and Instagram	In the highest accuracy of 93.7% followed by the KNN regression with 93.6%
Mahajan and Kaur (2021)	Social influence	A novel event recommendation system to suggest an event where the chances of a user's participation are high	IoT-CFR- IoT data	Better recommendation quality
Hodas et al. (2016)	The User's Personality Influences Content Engagement	An experiment combining electroencephalograms, personality surveys, and prompts	EEG Data	Personality and mood are highly correlated between friends via homophily
Bohacik et al. (2017)	Detecting Compromised Accounts	An anomaly model trained on the previous login data of users	Pokec	A real potential
Chaabani and Akaichi (2022)	Terrorist communities' evolution detection	-An Artificial Bee Colony optimization -BCTTC to track terrorist evolution	the Global Terrorism Database	Good results for small and large communities

Table 3 (continued)

Reference	Problem	Concept	Network	Proved
Aswani et al. (2017)	Identifying buzz in social media	-A hybrid artificial bee colony approach is integrated with k-nearest neighbors to identify and segregate buzz -A proposed hybrid bio-inspired approach	Twitter	Successfully giving an accuracy of 98.37%

to identify buzz in social media Aswani et al. (2017). This is also detected via a hybrid artificial bee colony approach integrated with k-nearest neighbors to identify and segregate buzz and a proposed hybrid bio-inspired approach.

6 Discussion and research challenges

One could wonder if we have adequately addressed the majority of the important issues underlying the identification of influencers, considering the volume of works that have been published in this area, such that the solutions suggested are sufficient for the majority of social network analysis tasks. Yet, in our opinion, a large variety of critical research issues should be solved until we can regard influencer identification as a problem-solving technique. We conclude that other significant fields have not yet been fully explored or studied, such as social network graphs and topology, and the relationship between influencers' identification and community detection, their networked audience switch MARWICK (2013) instead of studied fields, of which we have shown some relevant examples. Thus, we ponder unless we have successfully handled almost all of the critical problems related to this subject in a way that the solutions provided seem suitably sophisticated for most social network analysis tasks.

- Using social media, users first put their basic information, which can be false, into the social media guide to find their interesting topics. After they add friends and some aims or topics, social media algorithms expand their interests. Social media stars propose helpful suggestions for users to enrich their lives: friends from the same region; old school and faculty; and phone contact. It also suggests hobbies related to his work domain or studies. Then, after setting some preferences, social media stars suggest similar pages and friend types and eliminate ignored suggestions. Its suggestions are based on an intelligent system that analysts change needs, trembling all the time to suggest the most helpful. It certainly encourages users to use it more. The most important thing that users can do here is to search for new aims, new friends, or a new place and all the suggested updates. The challenge was that users could have a period or ephemeral aims. We think time has a big impact on social media users and the importance of a dynamic social media structure for researchers. There are many works in literature dealing with this idea, but no one has put into consideration the social media algorithms and strategies that determine social media influence.
- The majority of works in the literature have focused on the identification of the most influential people and significant nodes, called top-k. It represents an

interesting research area. Without considering Users can be influenced directly by their friends or groups of friends. They may influence them at the same time. Users share influence. All types of influencers have an important effect and can be the most influential because of their direct contact.

- Influence maximization is the trendiest topic that is being exploited and hopes to maximize users' influence like marketing publicity. But we do not have a search engine that deals with protecting users from negative and also positive users, who may likewise pose a risk by fostering wrath, resentment, hostility, hatred, and indignation, to minimize influence.
- User interests that change throughout time and trendy topics have a direct relationship with influence detection by simply comparing users' aims or interest topics. User interests change according to place or locality and season, influencing them, so we can predict influencers' node switches. For example, they can also change locales.
- Different works deal with the relationship strength and still cannot know the real life and direct relationship. Users may not communicate and discuss clear ideas on social media and may just exchange media or some symbols. Other contributions where the authors have aimed to exploit the users' locality information without considering that some users have double nationality or can aim with other country topics. The great challenge is, therefore, to find a solution to all these situations. The work of Sinha and Swearingen (2001) showed that suggestions from subject friends are better than recommender systems. The other challenge now is to study users' influencers' friends and predict their new aims.
- There are many types of influencers defined as: content creators MARWICK (2013) digital opinion leaders, moderators Enke and Borchers (2019) celebrity influencers, mega-influencers, and other works dealing with micro-influencers Campbell and Grimm (2019); Campbell and Farrell (2020); Leonardi and Monti (2020) that can be identified and found. So research is anticipated to rise.

7 Conclusion and prospects

7.1 Summary

Social media has emerged as a popular platform for members to discuss anything, give opinions, and express feelings and important moments about ideas and facts from everyday life Bao et al. (2013). In this article, we have provided an overview of the current state and future directions of influencer identification on social media. We have reviewed and classified the main existing contributions in the literature into the identification

influencers category. At first, we tried to talk more about this domain and alleged it in some works. Then, we cited some other classifications that deal with the same subject, and at least we have applied our classification proposition in some works according to our knowledge. We have proposed some new ideas that can change the orientation of searches on social media users' interest topics, and why not aim to influence new works?

Our study has some limitations. Most importantly, we did not aim at the structure and topology of social media. We have also noted that we identify significant nodes that influence other nodes in the influenced community and also detect them.

We also know that social media is virtual, but behind this virtual exists a real person whose acts, interactions, and shows are real. They are influenced by internal factors like friends and family, college classmates, and next-door neighbors, and external factors like life coaches, travel bloggers, influence, fashionistas, fitness models, and makeup artists. So we have detailed some important influencers' identification.

7.2 Prospects

In this state, to conclude our survey, we aimed to synthesize the main research works proposed in the literature, solve the problem of identifying influencers, and present the various algorithms, methods, methodologies, and proposed frameworks. It remains a topical problem, given the importance of the social media community in recent years, and it will continue in the future. Three new avenues may be outlined for this work's prospects. The first step is to undertake more complete comparison research to provide more information to scholars and particularly on how to find social media influencers. In the second part of the study, we will run more experiments on a multitude of other standard data sets to confirm the principal approaches provided in the literature's performance and robustness. The third part consists of proposing a new approach to finding new influencers on social networks by using other techniques and criteria based on the new networks' users' characteristics.

Declarations

Conflict of interest The authors have no affiliation with any organization with a direct or indirect financial interest in the subject matter discussed in the manuscript.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

- Adedoyin-Olowe M, Gaber MM, Stahl FT (2014) A survey of data mining techniques for social media analysis. *Data Mining and Digital Humanities*
- Adnan TMT, Islam MS, Papon TI, Nath S, Adnan MA (2022) Uacd: a local approach for identifying the most influential spreaders in twitter in a distributed environment. *Soc Netw Anal Min* 12:37
- Aggarwal C, Subbian K (2014) Evolutionary network analysis: a survey. *ACM Comput Survey* 47:1–36
- Alorainy W, Burnap P, Liu H, Williams M, Gjommoni L (2022) Disrupting networks of hate: characterising hateful networks and removing critical nodes. *Soc Netw Anal Min* 12:27
- Alp ZZ, Ögüdücü SG (2018) Identifying topical influencers on twitter based on user behavior and network topology. *Knowl-Based Syst* 141:211–221
- Alvari H, Shaabani E, Shakarian P (2018) Early identification of pathogenic social media accounts. in *IEEE Xplore*
- Alzanin SM, Azmi AM (2018) Detecting rumors in social media: a survey. *Procedia Comput Sci* 142:294–300
- Andreassen CS, Pallesen S, Griffiths MD (2016) The relationship between addictive use of social media, narcissism, and self-esteem: findings from a large national survey. *Addictive Behav* 64:287–293
- Arora A, Bansal S, Kandpal C, Aswani R, Dwivedi YK (2019) Measuring social media influencer index- insights from facebook, twitter and instagram. *J Retail Consum Services* 49:86–101
- Aswani R, Ghreha SP, Kar AK, Chandra S (2017) Identifying buzz in social media: a hybrid approach using artificial bee colony and k-nearest neighbors for outlier detection. *Soc Netw Anal Min* 7(1)
- Aswani R, Kar AK, Ilavarasan PV (2018) Detection of spammers in twitter marketing: a hybrid approach using social media analytics and bio inspired computing. *Inform Syst Frontiers* 20(3):515–530
- Azaouzi M, Mnasri W, Romdhane LB (2021) New trends in influence maximization models. *Comput Sci Rev* 40:100393
- Azaouzi M, Rhom D, Romdhane LB (2019) Community detection in large-scale social networks: state-of-the-art and future directions. *Soc Netw Anal Min* 9:23
- Bao H, Li Q, Liao SS, Song S, Gao H (2013) A new temporal and social pmf-based method to predict users' interests in micro-blogging. *Decision Support Syst* 55:698–709
- Bian J, Yang Y, Chua T-S (2014a) Predicting trending messages and diffusion participants in microblogging network. In 'Proceedings of the 37th international ACM SIGIR conference on Research and development in information retrieval'. p. 537–546
- Bian J, Yang Y, Chua T-S (2014b) Predicting trending messages and diffusion participants in microblogging network. In 'Proceedings of the 37th international ACM SIGIR conference on Research and development in information retrieval'. p. 537–546
- Bian R, Koh YS, Dobbie G, Divoli A (2019) Identifying top-k nodes in social networks: a survey. *ACM Comput Surveys* 52:1–33
- Bohacik J, Fuchs A, Benedikovic M (2017) Detecting compromised accounts on the pokec online social network. in *International Conference on Information and Digital Technologies (IDT)*
- Budak C, Kannan A, Agrawal R, Pedersen J (2014) Inferring user interests from microblogs
- Campbell C, Farrell JR (2020) More than meets the eye: The functional components underlying influencer marketing. *Business Horizons* 63(4):469–479
- Campbell C, Grimm PE (2019) The challenges native advertising poses: Exploring potential federal trade commission responses and identifying research needs. *J Public Policy Marketing* 38(1):110–123
- Cercel D, Trausan-Matu S (2014) Opinion propagation in online social networks: a survey. in: *Proceedings of the 4th acm international conference on web intelligence*. in 'Opinion propagation in online social networks: a survey. In: Proceedings of the 4th ACM International Conference on Web Intelligence'. pp. 5–8
- Cervellini P, Menezes AG, Mago VK (2016) Finding trendsetters on yelp dataset. in 'IEEE Symposium Series on Computational Intelligence (SSCI)'. pp. 1–7
- Chaabani Y, Akaichi J (2022) Bees colonies for terrorist communities evolution detection. *Soc Netw Anal Min* 12
- Chader A, Haddadou H, Hidouci W-K (2017) All friends are not equal: weight-aware egocentric network-based user profiling. In 'IEEE/ACS international conference on computer systems and applications'. pp. 482–488
- Chia K-C, Hsu C-C, Lin L-T, Tseng H H (2021) The identification of ideal social media influencers: Integrating the social capital, social exchange, and social learning theories. *J Electron Commerce Res* 22
- Cialdini RB (2001) *Influence: Science and practice* (4th ed.). allyn and bacon
- Cialdini RB, Goldstein NJ (2004) Social influence: compliance and conformity. *Annu Rev Psychol* 55(1):591–621
- Cialdini RB, NJ G (2002) The science and practice of persuasion. in 'Cornell Hotel Restaur Adm' p. 43 40–50
- Dang Q, Gao F, Zhou Y (2016) Early detection method for emerging topics based on dynamic bayesian networks in micro-blogging networks. *Expert Syst Appl* 57:285–295
- Ding Z, Wang H, Guo L, Qiao F, Cao J, Shen D (2015) Finding influential users and popular contents on twitter. In 'International conference on web information systems engineering'. pp. 267–275
- Domingos P, Richardson M (2001) Mining the network value of customers. In 'Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining'. p. 57–66
- Elanor C (2013) Csr communication strategies for organizational legitimacy in social media. *Corporate Communicat* 18(2):228–248
- Enke N, Borchers NS (2019) Social media influencers in strategic communication: A conceptual framework for strategic social media influencer communication. *Inter J Strategic Communicat* 13(4):261–277
- Erlandsson F, Bródka P, Borg A, Johnson H (2016) Finding influential users in social media using association rule learning. *Entropy* 18:164
- Fang Q, Sang J, Xu C, Rui Y (2014) Topic-sensitive influencer mining in interest-based social media networks via hypergraph learning. *IEEE Transct Multi* 16:796–812
- Han X, Wang L, Crespi N, Park S, Cuevas A (2015) Alike people, alike interests? inferring interest similarity in online social networks. *Decis Support Syst* 69:92–106
- Harrigan P, Daly TM, Coussement K, Lee JA, Soutar GN, Evers U (2021) Identifying influencers on social media. *Inter J Inform Manage* 56:102246
- Helali L, Omri MN (2021) A survey of data center consolidation in cloud computing systems. *Comput Sci Rev* 39:100366
- Hernandez M, Hildrum K, Jain P, Wagle R (2013) Constructing consumer profiles from social media data. In 'IEEE international conference on big data'. pp. 710–716
- Hodas NO, Butner R, Corley C (2016) How a user's personality influences content engagement in social media. *Social Informatics*. p. 481–493
- Imamori D, Tajima K (2016) Predicting popularity of twitter accounts through the discovery of link-propagating early adopters. in 'CIKM'
- Iraklis M, Williams HT (2020) Good and bad events: combining network-based event detection with sentiment analysis. *Soc Netw Anal Min* 10(1):1–12

- Izdihardian WA, Ruldeviyani Y (2021) Detecting social media influencers of airline services through social network analysis on twitter: A case study of the Indonesian airline industry. In '3rd East Indonesia Conference on Computer and Information Technology (EIConCIT). IEEE'
- Jain S, Sinha A (2020) Identification of influential users on twitter: A novel weighted correlated influence measure for covid-19. *Chaos Solitons and Fractals*. 139:110037
- Jia J, Wang B, Zhang L, Gong NZ (2017) Attriinfer: Inferring user attributes in online social networks using markov random fields. In 'Proceedings of the 26th International Conference on World Wide Web'. p. 1561–1569
- Jin SV, Muqaddam A, Ryu E (2019) Instafamous and social media influencer marketing. *Market Intell Plan Emerald Publishing Limit* 37:567–579
- Jin SV, Ryu E (2020) "i'll buy what she's wearing": The roles of envy toward and parasocial interaction with influencers in instagram celebrity-based brand endorsement and social commerce. *J Retail Consumer Service* 55:102121
- Kaple M, Kulkarni K, Potika K (2017) Viral marketing for smart cities: Influencers in social network communities. In 'IEEE Third International Conference on Big Data Computing Service and Applications (BigDataService)'. pp. 106 – 111
- Karamshuk D, Lokot T, Prymak O, Sastry N (2016) Identifying partisan slant in news articles and twitter during political crises. in 'International Conference on Social Informatics (SocInfo)'
- Kempe D, Kleinberg J, Tardos E (2003) Maximizing the spread of influence through a social network. In 'Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining'. p. 137–146
- Kitchenham B, Charters S (2007) Guidelines for performing systematic literature reviews in software engineering (version 2.3). EBSE Technical Report
- Leonardi S, Monti D (2020) Mining micro-influencers from social media posts. In 'Proceedings of the 35th Annual ACM Symposium on Applied Computing'
- Li R, Wang S, Deng H, Wang R, Chang KC-C (2012) Towards social user profiling: Unified and discriminative influence model for inferring home locations. In 'Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining'. p. 1023–1031
- Liselot Hudders, De Jans Steffi, DV M (2021) The commercialization of social media stars: A literature review and conceptual framework on the strategic use of social media influencers. *International Journal of Advertising*. 40(3):327–375
- Mabrouk O, Hlaoua L, Omri MN (2018) Fuzzy twin svm based-profile categorization approach. In '14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)'. pp. 547 – 553
- Mabrouk O, Hlaoua L, Omri MN (2020) Exploiting ontology information in fuzzy svm social media profile classification'', applied intelligence. *Appl Intell* 51:3757–3774
- Mahajan P, Kaur PD (2021) Harnessing user's social influence and iot data for personalized event recommendation in event-based social networks. *Soc NetW Anal Min*. 11(1)
- Mahyar H, Hasheminezhad R, Ghalebi E, Ali Nazemian RG, Movaghar A, Rabiee HR (2018) Identifying central nodes for information flow in social networks using compressive sensing. *Social Network Analysis and Mining*
- MARWICK AE (2013) Status update: Celebrity, publicity, and branding in the social media age. yale university press
- More JS, Lingam C (2019) A gradient-based methodology for optimizing time for influence diffusion in social networks. *Soc Netw Anal Min*. 9(1)
- Ognyanova K (2017) Multistep flow of communication: Network effects. in 'New York NY: Wiley-Blackwell'. pp. 1–10
- Oren S (2019) Algorithmic personalization and the two-step flow of communication. *Communication Theory*
- Ouni S, Fkih F, Omri MN (2022a) Bert and cnn based tobeat approach for unwelcome tweets detection. *Social Network Analysis and Mining*. 12
- Ouni S, Fkih F, Omri MN (2022b) Bots and gender detection on twitter using stylistic features. In '14th International Conference on Computational Collective Intelligence'
- Ouni S, Fkih F, Omri MN (2022c) Novel semantic and statistic features-based author profiling approach. *J Ambient Intell Human Comput*
- Peng S, Wang G, Xie D (2017) Social influence analysis in social networking big data: opportunities and challenges. *IEEE Netw* 31:11–17
- Peng S, Wang G, Zhou C, Wang C, Yu S, Niu J (2017) An immunization framework for social networks through big data based influence modeling. *IEEE Transact Dependable Secure Comput* 16:984–995
- Peng S, Zhou Y, Caoc L, Yud S, Niue J, Jiaf W (2018) Influence analysis in social networks: a survey. *Netw Comput Appl* 106:17–32
- Piao G, Breslin JG (2018) Inferring user interests in microblogging social networks: a survey. *User Modeling User-Adapted Interact* 28:277–329
- Plantie M, Crampes M (2013) Survey on social community detection. *Computer Communications and Networks*, Springer-Verlag, London, Social Media Retrieval
- Probst F (2013) Who will lead and who will follow: identifying influential users in online social networks. *Business Inform Syst Eng* 5:179–193
- Pudjajana AM, Manongga D, Iriani A, Purnomo HD (2018) Identification of influencers in social media using social network analysis (sna). In 'International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)'. pp. 400 – 404
- Qasem Z, Jansen M, Hecking T, Hoppe H (2015) On the detection of influential actors in social media. In 'th International Conference on Signal-Image Technology and Internet-Based Systems (SITIS)'. pp. 421–427
- Qasem Z, Jansen M, Hecking T, Hoppe H (2017) Influential actors detection using attractiveness model in social media networks. *Inter Workshop Complex Netw Appl* 693:123–134
- Rodríguez-Vidal J, Gonzalo J, Plaza L, Sánchez HA (2019) Automatic detection of influencers in social networks: authority versus domain signals. *J Associat Inform sci technol* 70:675–684
- Saez-Trumper D, Comarela G, Almeida V, Baeza-Yates R, Benevenuto F (2012) Finding trendsetters in information networks. In 'Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining'. p. 1014–1022
- Sendi M, Omri MN (2015) Biomedical concept extraction based information retrieval model: application on the mesh. In 'International Conference on Intelligent Systems Design and Applications (ISDA)', 14–16 December 2015. pp. 40–45
- Sendi M, Omri MN, Abed M (2017) Possibilistic interest discovery from uncertain information in social networks. *Intell Data Anal* 21:1425–1442
- Sendi M, Omri MN, Abed M (2019) Discovery and tracking of temporal topics of interest based on belief-function and aging theories. *J Ambient Intell Human Comput* 10:3409–3425
- Shah B, Verma AP, Tiwari S (2018) User interest modeling from social media network graph, enriched with semantic web. In 'Proceedings of International Conference on Computational Intelligence and Data Engineering'. pp. 55–64
- Shahr HSA, Yazdani S, Afshar L (2019) Professional socialization: an analytical definition. *J Medical Ethics History of Med*. pp. 12–17

- Sheikhahmadi A, Nematbakhsh MA, Zareie A (2017) Identification of influential users by neighbors in online social networks. *Physica A: Statist Mech Appl* 486:517–534
- Singlaand P, Richardson M (2008) Yes, there is a correlation: - from social networks to personal behavior on the web. in 'In Proceeding of the 17th international conference on World Wide Web (WWW'08)'. p. 655–664
- Sinha R, Swearingen K (2001) Comparing recommendations made by online systems and friends. *DELOS Personalisation and Recommender Systems in Digital Libraries*
- Souden I, Omri MN, Brahmi Z (2022) A survey of outlier detection in high dimensional data streams. *Comput Sci Rev* 44:100463
- Staab S, Domingos P, Mika P, Golbeck J, Ding L, Finin T, Joshi A, Nowak A, Vallacher R (2005) Social networks applied. *IEEE Intell Syst* 20:80–93
- Subbian K, Sharma D, Wen Z, Srivastava J (2014) Finding influencers in networks using social capital. *Soc Netw Anal Min* 11:219
- Sun B, Ng VT (2012) Identifying influential users by their postings in social networks. *Inter Workshop Model Soc Med* 8329:128–151
- Sun J, Tang J (2011) A survey of models and algorithms for social influence analysis. in 'Social Network Data Analytics'. pp. 177–214
- Sun Q, Wang N, Zhou Y, Luo Z (2016) Identification of influential online social network users based on multi-features. *Inter J Pattern Recognition Artif Intell* 30:1659015
- Sunil MJ, Lingam C (2019) A si model for social media influencer maximization. *Appl Comput Inform* 15(2):102–108
- Tabassum S, Pereira FSF, Fernandes S, Gama J (2018) Social network analysis: an overview. *Wiley Interdisciplinary Rev: Data Mining and Knowledge Discovery*. 8(5):e1256
- Tsugawa S, Kimura K (2018) Identifying influencers from sampled social networks. *Physica A*. 507:294–303
- Utz S (2010) Show me your friends and i will tell you what type of person you are: How one's profile, number of friends, and type of friends influence impression formation on social network sites. *Comput-Mediated Communicat* 15:314–335
- Wang H, Huang X, Li L (2018) Microblog oriented interest extraction with both content and network structure. *Intell Data Anal* 22:515–532
- Wang Z, Liu H, Liu W, Wang S (2020) Understanding the power of opinion leaders' influence on the diffusion process of popular mobile games: Travel frog on sina weibo. *Comput Human Behavior*. 109:106354
- Xiang R, Neville J, Rogati M (2010) Modeling relationship strength in online social networks. In 'Proceedings of the 19th international conference on World wide web'. p. 981–990
- Xie H, Li Q, Mao X, Li X, Cai Y, Rao Y (2014) Community-aware user profile enrichment in folksonomy. *Neural Netw* 58:111–121
- Yang Y, Zhang C, Fan C, Yao W, Huang R, Mostafavi A (2019) Exploring the emergence of influential users on social media during natural disasters. *Inter J Disaster Risk Reduct* 38:101204
- Yang Z, Wilison C, Wang X, Gao T, Zhao BY, Dai Y (2014) Uncovering social network sybils in the wild. *ACM Transact on Knowl Discovery from Data* 8(1):1–29
- Zareie A, Sheikhahmadi A, Jalili M (2019) Identification of influential users in social networks based on users' interest. *Inform Sci* 493:217–231
- Zarrinkalam F, Kahani M, Bagheri E (2017) Mining user interests over active topics on social networks. *Inform Process Manage* 54:339–357
- Zarrinkalam F, Kahani M, Bagheri E (2019) User interest prediction over future unobserved topics on social networks. *Inform Retrieval J* 22:93–128
- Zhang J, Tomonaga S, Nakajima S, Inagaki Y, Nakamoto R (2015) Finding prophets in the blogosphere: Bloggers who predicted buzzwords before they become popular. In 'Proceedings of the 17th International Conference on Information Integration and Web-based Applications and Services'. p. 1–10
- Zhao Z, Zhou H, Zhang B, Ji F, Li C (2019) Identifying high influential users in social media by analyzing users' behaviors. *J Intelligent Fuzzy Syst* 36:6207–6218
- Zheng C, Zhang Q, Young SD, Wang W (2020) On-demand influencer discovery on social media. In 'Proceedings of the 29th ACM international conference on information and knowledge management'
- Zhou F, Lü L, Mariani MS (2019) Fast influencers in complex networks. *Communicat Nonlinear Sci Num Simulat* 74:69–83

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.