

A Review of Methods for Detecting Rumors on Social Networks Using Social Circle Mining

Alaa Alassaf^{1*}, Delel Rhouma^{1,2}

¹Department of Computer Science, College of Computer, Qassim University, Buraydah, Saudi Arabia

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

Email: *a.alassaf@qu.edu.sa, d.rhouma@qu.edu.sa

How to cite this paper: Alassaf, A. and Rhouma, D. (2025) A Review of Methods for Detecting Rumors on Social Networks Using Social Circle Mining. *Social Networking*, 14, 15-44.

<https://doi.org/10.4236/sn.2025.142002>

Received: March 21, 2025

Accepted: April 27, 2025

Published: April 30, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Currently, social platforms provide places for people not only to gather information but also to generate and propagate rumors. Consequently, rumor detection has become a major global task for mining fake information from social networks. Although social circles have the capacity to describe users' behavior preferences and to impact the scope and spreading speed of rumors, numerous studies have ignored them when designing rumor prediction models. To address this oversight, we conducted a technical investigation and comparison of state-of-the-art procedures for detecting rumors, focusing on the role of data mining in social circles. This survey will assist researchers in determining the most effective techniques and appropriate future research directions.

Keywords

Rumor Detection, Social Network Analysis, Social Circles, Graph Mining

1. Introduction

Our lives and interactions with others are significantly affected by social media today. In the past, people relied on voice transmitters such as mongers and drums to communicate, limiting the spread of news to just a few kilometers. Subsequently, people turned to print media, such as newspapers and magazines, and later to radio and television. However, a barrier between journalists and the public prevents the latter from sharing their opinions and ideas on a wide scale. Then came social media, which expanded the scope of knowledge and participation for exchange among participants. Through social media platforms such as Facebook and X (formerly Twitter), in a few seconds, information published by one person can reach millions of individuals [1].

The era of social media began two decades ago and has continuously gained popularity. Platforms such as LinkedIn, MySpace, Facebook, YouTube, X, Google+, Instagram, WhatsApp, and Snapchat caused a major revolution in our communication and social life [2]. The ease and flexibility of social media platforms, especially X, where anyone can share and collect information about news regardless of its veracity, have made them fertile ground for rumor mongers to spread misleading information that may prompt disturbances and unexpected responses [3]. Owing to the unorganized nature of these platforms, rumors spread widely. Therefore, studying and analyzing rumors on social media platforms is a crucial topic in the field of social media analysis [4].

Data are continuously generated by social networks every minute. Studies have demonstrated that Facebook users share 684,000 pieces of material in a single minute, while X users send more than 100,000 tweets. Hand-listing rumors from such a massive dataset is not feasible [5]. Numerous studies have analyzed the problem of rumors spreading on social networks using various techniques such as machine learning and deep learning. However, they have often neglected the key role of users in determining the spread of rumors. Even among studies investigating rumor sources, few have examined the impact of social media mining on rumor detection [6]. Therefore, this study seeks to analyze the techniques used in rumor detection, focusing on social circle mining in rumor detection and the associated challenges. The main contributions of this study are as follows.

- We reviewed the methods used for rumor detection and highlighted the importance of implementing a system that combines deep learning models and social circle mining.
- We comprehensively reviewed and classified rumor detection methods into different types based on the underlying approach.
- We compare and analyze rumor detection methods based on the approach used in rumor detection from several perspectives.
- We identify the prevailing research challenges in rumor detection and propose practical solutions to these issues.

In this study, we used a systematic literature review (SLR) methodology [7] to examine the basic concepts related to rumor detection on social media. We also investigated several methods and approaches designed to detect rumors on social media platforms, focusing on the role of social circle mining and the challenges facing this research area.

The remainder of this paper is organized as follows: In the second section, we review the most important definitions of rumors from previous studies. In the third section, we divided the methods used for rumor detection into four categories. In the fourth section, we present the concepts of social media, rumors, and social circle mining for rumor detection. The technical details of the various models are compared in the fifth section. In the sixth section, the challenges related to rumor detection in social networks are discussed. The seventh section concludes the study.

2. Rumor Definition

There are two common ways to describe a rumor. Rumors are often understood to be “distorted, exaggerated, irrational, and inauthentic information” [8]. In academic studies, rumors are usually described as messages that have not been verified or validated, such as “an unverified proposition that is in general circulation” [9] or “an unverified message passed from one person to another that refers to an object, person, or situation” [10], including rumors that are subsequently shown to be untrue as well as those later confirmed as correct [11]. “Unverified and instrumentally relevant information statements in circulation” [12] are considered rumors, regardless of whether this unconfirmed information is later confirmed as wholly untrue, somewhat true, or neither. Another definition describes a rumor within the context of breaking news stories as a “circulating story of questionable veracity, which is credible though hard to verify, and produces sufficient skepticism and/or anxiety to motivate finding out the actual truth” [13].

This study embraces the standard academic definition of a rumor, which states that a rumor is a message that has not yet been verified by the recipient and is unreliable in terms of accuracy [14]. Rumors in this study are considered widely circulated statements or pieces of information that have not been verified as true or false, which are often spread by word of mouth or through social media and lead to misleading information or social confusion if not verified.

3. Literature Review of Rumor Detection Methods

A contemporary tale or report with questionable or ambiguous veracity may be considered a rumor. A rumor may also be described as an association that begins with one or more sources and develops over time. The veracity of any given occurrence is determined via rumor detection. Research on developing rumor detection and prediction systems has gained considerable popularity following increased reliance on social media for news and important information. When performing rumor detection, several issues must be addressed, such as the variety of data, identifying the most recent rumors from historical data, and determining the source of the rumor. A rumor classification system generally consists of four components: situation classification, truth classification, tracking, and detection. When building rumor classification systems, temporal characteristics must be determined because rumors usually appear with breaking news that the system has not previously encountered. Therefore, the system must be able to deal with new, unexpected rumors and detect them automatically and immediately [4].

Related to rumor detection, rumor prediction first appeared in 2017, with the goal of predicting the possibility that a message will become a rumor at the time of publication rather than—as with rumor detection—merely identifying rumors after they occur. In determining whether a message or information may become a rumor in the future, rumor prediction requires quick and immediate decisions before the rumor spreads and places people at risk. Prediction systems may face difficulties related to short sentences and the volume of data flow. However, it is

important to identify rumors as soon as possible before the process becomes overly complicated [5]. **Figure 1** displays the steps in detecting rumors, classifying stances, and predicting truth.

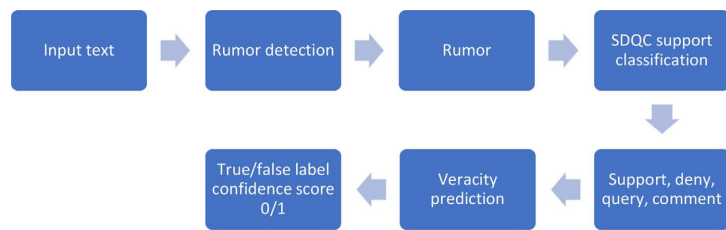


Figure 1. Flowchart of the rumor detection and classification process.

The spread of rumors negatively impacts people's lives. Therefore, obtaining reliable information during crises is of utmost importance. The current strategies for rooting out rumors on social media platforms incorporate a series of tools and methods that identify, trace, and suppress misinformation through various computational, statistical, and machine-learning techniques to distinguish between facts and rumors in the information posted on social media [15]. In the following section, we outline the most recent and important approaches used in previous studies for detecting rumors.

3.1. Machine Learning-Based Approaches

Machine learning (ML) is a branch of artificial intelligence (AI). Using algorithms learned via a large dataset, ML builds models with the ability to perform tasks without explicit learning, ranging from image classification to future prediction procedures [16]. Lying at the intersection of several scientific disciplines, ML is a rapidly growing field encompassing computer science, statistics, data science, and AI. ML has made significant strides with the advancement of algorithms and the abundance of data available through the Internet [17]. Rumor detection using ML frameworks is an exciting and rapidly developing research area. The ML process flow is depicted in **Figure 2**, where the best characteristics are taken from the input and then categorized in a binary classification system as rumors (R) or non-rumors (NR).

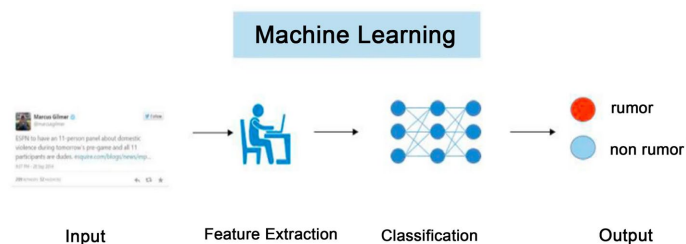


Figure 2. Use of machine learning for rumor detection.

3.1.1. Supervised Approaches Units

In a supervised learning approach, models classify posts as rumors or not by train-

ing on a labeled dataset. This can be accomplished using several algorithms, including logistic regression, decision trees, the Naive Bayes algorithm, and support vector machines (SVM) [18]. Rumors can be accurately detected by filtering and analyzing language features using the supervised ML framework bifold [19].

1) Methodology of Supervised Approaches

A substantial body of research on supervised approaches can be traced back to Castillo *et al.*'s [20] foundational study on assessing the reliability of information disseminated through social media networks examination of the reliability of information distributed through social media networks. Castillo *et al.* [20], who were the first to develop features for evaluating credibility on Twitter [21], based their research on the notion that users can assess the reliability of information using specific cues found in social media environments.

Castillo *et al.* created a dataset to investigate Twitter's trustworthiness. A team of human assessors then assigned a label to each issue based on two criteria: NEWS, which provides a fact that may be of interest to others, and CHAT, which is a message based on the author's personal ideas and/or chats with friends. Another panel of judges rated each item's trustworthiness within these categories [21]. Subsequently, a collection of attributes was extracted to construct a classifier that automatically ascertains whether a subject is associated with noteworthy data or occurrences and, if so, automatically evaluates the subject's degree of believability. Four categories of features were distinguished: propagation-based, message-based, user-based, and topic-based features [21].

The best-first selection approach was used to select features that contributed most to the task. Sampling with replacement was used to select a random sample from the training dataset, and a uniform distribution was used to determine the likelihood of obtaining an instance from each of the three classes (NEWS, CHAT, and UNSURE). A variety of supervised learning techniques were used [21].

2) Nonparametric Machine Learning Models

The term nonparametric refers to models in which the underlying data distribution is not assumed to have a fixed shape or a fixed number of parameters. Instead, they adjust to the structure of the data, which can be used with both supervised and unsupervised ML models. In classifying fake news according to its prevalence level, in addition to the use of textual information (text and headlines) and content features, preventive strategies can also be implemented to create early predictions regarding the spread of fake news [22].

3) Logistic Regression Model

Logistic regression models are statistical models used to estimate the probability of a binary event, such as yes/no, true/false, and, in our context, common/uncommon, using one or more independent variables [23]. Logistic regression has been used in the health sector to classify health-related rumors. According to a study on rumor veracity categorization, the length of rumor headlines or statements and the use of pictures in rumor statements were adversely connected with the chance that a rumor was real. The following characteristics were positively

connected with the likelihood that a rumor would be true: hyperlinks, the names of people or places mentioned in the rumor, and hints about the source of information [24].

Using regression tasks, fact-checking platforms are developing computational methods to predict the popularity of false rumors, focusing on immediate contextual information, such as rumor content and user profile information [25].

4) Support Vector Machines (SVMs)

In the 1950s, the SVM method first appeared, subsequently developing as a kernel computer [16]. The SVM is a supervised ML model that is applied to classification and regression problems and works by ideally dividing data points into different classes to determine the ideal hyperplane in the feature space [26].

In this study, Qin *et al.* used an SVM on a training set and based rumor prediction on a fixed threshold technique to maximize prediction accuracy. Furthermore, they used pseudo-feedback, novelty-based, and content-based features. The proposed method can reliably forecast whether an item eventually becomes a rumor, enabling the refutation of false information before it spreads and causes harm to society. Rumor prediction significantly increases the precision of cutting-edge rumor detection software [5].

5) Random Forest (RF)

The RF model, which is also widely used for categorization, uses several decision trees. In this model, each decision tree serves as a classifier, and a random sampling algorithm aggregates all classification results before the final classification decision is reached. RF has been demonstrated to be a highly accurate and reliable algorithm compared with various rumor-refuting microblogs [27]. However, it has not been proven to be as effective as XGBoost in identifying users who are more likely to deny rumors [28].

6) Extreme Gradient Boosting (XGBoost)

XGBoost is a recently developed algorithm that is increasingly popular in data analysis and industrial circles. In contrast to RF, it trains trees in a serial and interactive manner, as opposed to a parallel and autonomous manner, using a different type of tree called the classification and regression tree (CART) [29].

XGBoost was used in a study to determine which users were more inclined to deny rumors. The prediction model was enhanced by text analysis using natural language processing (NLP), which also demonstrated high generalizability and the absence of any significant group behavior characteristics in the retweeted prediction. A valid and robust distinguishing XGBoost model was then trained and validated to predict who would retweet disaster-related rumor-refuting microblogs [28].

The XGBoost model has been used, along with several content-, user-, and topic-based features, to automatically detect rumors in Arabic [30]. In another study, an XGBoost classification model was trained utilizing NLP, together with other user characteristics, such as user sentiment and current interests, allowing it to examine the relationship between user characteristics and rumor-disputing be-

havior in five major rumor categories—economics, society, politics, disasters, and the military [31].

3.1.2. Unsupervised-Based Approaches

The unsupervised learning approach involves training models on unlabeled data with the goal of uncovering underlying structures or hidden patterns in the data. Illustrative techniques encompass anomaly detection and dimensionality reduction (principal component analysis, t-small number ensemble), as well as clustering (K-means, linear clustering) [16] [32].

1) Feature Analysis

Feature analysis is an unsupervised learning method that aims to identify hidden patterns or structures in data. In our context, many features have been examined for rumor detection, such as retweet rate and word distribution characteristics. A study concluded that the retweet rate was insufficient to distinguish rumors from other items. In contrast, the differences in word distribution may be an important indicator for classifying rumors [33].

2) Clustering

Clustering is a method that uses models to group data points based on the similarity of the data. It seeks to identify groups in a dataset that occur naturally without labels. Examples of clustering algorithms include K-means, hierarchical clustering, and DBSCAN [34].

A study on clustering Twitter political rumor detection used five structural and chronological characteristics. In the first phase, tweets with consistent address links were grouped. In contrast, in phase two, similar clusters debating continuous news were combined into a single cluster with similar circular functions [35]. Using clustering to group similar tweets and semantic and sentiment analyses to identify differences in information supplied by various sources, these techniques work together to help uncover rumors on Twitter [36].

3.1.3. Self-Supervised-Based Approaches

This method uses unlabeled data by generating supplementary tasks to enhance the primary supervised task. To better identify and detect rumors, the self-supervised rumor detection (SRD) framework employs contrastive learning on heterogeneous information sources and is part of the unsupervised learning subcategory. This technique does not rely on labels from other sources but instead on the data itself to generate supervisory signals. The applications used in this framework include NLP, computer vision, and speech recognition. Common techniques include removing the signal from the image, which includes filling in the missing parts of the image, contrastive learning, and word prediction using language models such as BERT (bidirectional encoder representations from transformers) [37].

1) Contrastive Learning

By improving the model's capacity to discern between accurate and inaccurate information, contrastive learning has the potential to be a valuable tool for identifying rumors. It has been used in conjunction with other technologies, which

enhanced performance [38].

2) BERT Models

Researchers have demonstrated that using BERT in different ways significantly enhances the accuracy of rumor detection compared with traditional methods [39] [40]. Such increases are owing to BERT's enhanced capacity to comprehend the subtleties and contextual meaning of language used in rumors.

3.2. Deep Learning-Based Approaches

Deep learning is a subfield of ML that employs multilayered artificial neural networks (ANNs), popularly known as “deep” neural networks (DNNs), to represent complex patterns in data. Neural networks can acquire knowledge from vast datasets because of their inherent design, which emulates the cognitive processes of the human brain. Deep learning works well for applications such as audio and picture recognition and NLP [41] [42] and has demonstrated to be a clear winner over traditional classifiers in a variety of ML applications, including speech recognition, object identification, and sentiment categorization. The goal of DNN-based techniques is to automatically develop deep representations of rumor data [43]. The use of deep learning for rumor detection is illustrated in **Figure 3**.

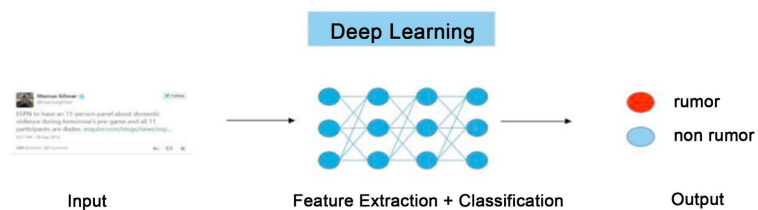


Figure 3. Deep learning for rumor detection.

Several deep learning methods are employed for rumor detection on social media. The subsections below outline some common approaches.

3.2.1. Recurrent Neural Network (RNN)-Based Approaches

A diverse class of feedforward neural networks, RNNs are mostly employed to handle variable-length sequential or time series data. RNNs are used in a variety of applications, such as categorization and sequence creation [3]. Given their structure, RNNs form models of rumor data as sequential data, capturing the dynamic temporal signals characteristic of rumor diffusion through direct cycle connections between units [43]. For instance, integrating an RNN model with a Word2Vec model provides a novel approach to rumor detection. This hybrid model comprises two primary components: Word2Vec and long short-term memory (LSTM)-RNN. A common problem faced in detecting rumors in breaking news is overlapping topics and out-of-vocabulary (OOV) items; parallel training of components helps mitigate such issues, providing a powerful solution for detecting rumors in real time. An RNN-based approach was used to leverage both the propagation patterns and linguistic content of posts for early rumor detection, allowing

the RNN to learn the relationships between posts and their content [44].

The CSI model, which refers to the components that make up the model: Capture, Score, and Integrate, is an innovative rumor detection technique that combines text, user comments, and promotions. The use of an RNN enables the understanding of the dynamics involved in the spread of rumors. The hybrid CSI model enhances accuracy and reliability by integrating several data profiles. Furthermore, it offers valuable insights for further studies by generating concealed user and object representation. The CSI model represents a significant advancement in rumor detection. Providing a comprehensive reaction to misinformation shared on social media enables swift intervention and mitigation [45].

3.2.2. Convolutional Neural Network (CNN)-Based Approaches

A CNN utilizes a structure that models significant semantic features, thereby enabling it to extract local and global features and reveal high-level interactions [43]. Recent research has revealed constraints in using RNNs for the early identification of rumors because they rely heavily on recent input data. A solution to this problem was the development of a convolutional approach for misinformation identification (CAMI). The proposed approach uses doc2vec to represent clusters of microblog entries and a two-layer CNN for classification, enabling the efficient extraction of crucial information and high performance [46].

In another study, a novel event adversarial neural network (EANN) was developed specifically for detecting multimodal false news, particularly for recent events. The EANN incorporates modules for event discrimination, bogus news identification, and multimodal feature extraction. It incorporates CNNs to extract characteristics from both textual and visual data [47].

3.2.3. Graph Neural Network (GNN)-Based Approaches

These neural networks are designed to operate with graph-structured data. GNNs are especially helpful for tasks in which relationships and interactions between entities are crucial. Among its most important applications, social network analysis is used to detect communities, predict user behavior, and identify influential users, with other applications including recommendation systems and NLP. The global structural characteristics from graphs or trees are better captured by a graph convolutional network (GCN) than by the aforementioned deep learning models. Driven by CNN's achievements of CNNs in computer vision, GCN have proven their performance at the forefront of several tasks using graph data. Several recent studies have used GCNs to detect and predict rumors on social media networks [48].

Deep learning methods often overlook the structures of wide dispersions in rumor detection. The bi-directional graph convolutional network (Bi-GCN) model explores both the propagation and dispersion characteristics of rumors using a top-down directed graph and an opposite graph of rumor diffusion. The proposed model has been proven to be effective in early rumor detection on Twitter [49].

A gated GNN-based algorithm [50] called propagation graph neural network

(PGNN) can generate powerful representations for each node in the graph. The global embedding with PGNN (GLO-PGNN) and ensemble learning with PGNN (ENS-PGNN) models adopt different classification strategies and include attention mechanisms to dynamically adjust the node weights to further improve performance.

The use of a unified GNN model and a community-enhanced vulnerability propagation method can improve rumor prediction and identify user vulnerabilities [51]. A gated GNN was used to learn semantic dependencies across segments in a graph-based pivotal semantic mining framework [52]. Given the many shortcomings of previous techniques and the need for improved reliability and accuracy of detection mechanisms, these sophisticated models and methodologies mark a substantial advancement in the field of rumor and false news identification.

3.3. Hybrid Approaches

Hybrid models combine more than one approach to increase the accuracy and effectiveness of the model. In the context of detecting rumors, hybridization often combines traditional ML methods with deep learning to enhance the effectiveness of the model and its ability to detect rumors [4].

In rumor detection, hybrid methods can be classified as follows:

3.3.1. Combinations of Features

This model combines several features that enable it to predict rumors with higher accuracy. The combination of features such as text, user behavior, and network structure increase the model's effectiveness and accuracy [38] [40] [53] [54].

3.3.2. Ensemble Learning

The ensemble learning approach can combine the capabilities of more than one model in prediction to improve and increase efficiency, which makes it superior to single models in terms of their accuracy and flexibility. Examples of common methods for this approach include boosting, stacking, and bagging [38].

3.3.3. Multimodal Approaches

Multimodal approaches employ several features together to offer a comprehensive view of a situation. In predictions, these approaches can use a combination of text, images, and audio to enhance model functioning [55]-[57].

3.3.4. Hybrid Algorithms

More than one algorithm can be exploited to explore the benefits that these methods provide since most ML algorithms, and many neural network structures can be complemented with deep learning methods for higher levels of accuracy and enhanced analysis [39] [54] [56] [58] [59].

3.3.5. Integration of Deep Learning

This method uses several deep learning techniques, like CNNs and RNNs, to cre-

ate a more accurate and useful model or way of predicting or extracting features that make use of the strengths and weaknesses of many neural network designs [38] [40] [55] [56] [59].

3.4. Network-Based Approaches

For rumor detection, network-based approaches have been proposed that focus on the structure of social networks to identify how information and rumors are distributed and spread across them. Community discovery algorithms, centrality measures, and influence modeling often aid in the identification of influential nodes or groups within a community, which is a major factor in the spread of rumors across the network. By tracking the information spread and rumor pathways in hypothetical networks, network-based approaches expect to forecast, observe, and eventually control the harm and negative consequences of rumors [21].

The reliability analysis of rumor propagation modeling is currently one of the most well-known multidisciplinary study areas regarding this topic. Existing studies have primarily focused on three areas:

The enhancement of the classical rumor propagation model based on factors, such as, heterogeneous user representation, was proposed for automatically distinguishing rumors from credible messages posted on social media [21]. The analysis of rumor propagation on online social platforms through the integration of other disciplinary theories, such as game theory, where social anxiety is considered, the intensity of rumors, and public perceptions of rumors and anti-rumors are considered important factors [60]. The combination of more than one technology that focuses on prominent nodes to reduce the propagation of rumors in online social networks has been shown to increase the effectiveness of control mechanisms when high-centrality users are the focus of attention [61].

Social networks enable individuals to participate in the exchange of information, providing large amounts of diverse data. It is worth noting that conventional techniques for social circle finding, similar to community detection, use clustering through statistical inference from a network topology perspective. In contrast, global mining requires considerable resources due to the size of social networks. Because of this difficulty, many academics have concentrated on mining the social circles of users' egocentric networks. Most social circle-finding algorithms are only applicable to specific datasets, making it difficult to work directly with them to detect rumors [6] [62].

3.5. Transformer-Based Architectures

Following the remarkable success of transformer-based embeddings in natural language tasks, researchers have become interested in applying them to classify rumors on social media platforms. Unlike traditional word embedding methods, the former excels at capturing the contextual meaning of a text by examining words to the left and right, producing textual representations that are ideal for detecting rumors [63]. A novel approach, called the 'Memory-Enhanced Transformer with

Graph Convolutional Networks' (GCNs-MT), was presented for detecting rumors on social media platforms [64]. By combining GCNs with long-short-term memory cells and multi-head attention mechanisms in transformers to capture local and global dependencies in rumor propagation, we have proven our ability and efficiency in detecting rumors, addressing the challenges posed by the rapid spread of rumors on social media platforms.

As fake news and rumors spread rapidly, it is important to develop a model for automatic rumor detection. The transformer-based multimodal interactive fusion (TMIF) model [65] relies on a transformer to map multimodal feature representations to the same data domain for fusion purposes. It can capture multilevel dependencies between multimodal content while minimizing the impact of multimodal heterogeneity. Furthermore, it demonstrated superior performance in the early detection of rumors.

4. Social Media

Social media (SM) is an umbrella term for various online platforms, including blogs, social networks, and virtual worlds. The phrase "social media" was originally used in 1994 by the Matisse online media environment in Tokyo, Japan. Initially, SM platforms were created and introduced during the preliminary stages of commercial Internet. SM, now recognized as one of the most significant applications of the Internet, has substantially expanded in terms of platform diversity and user engagement over time. SM has many uses, not just sharing photos and promoting and advertising [66].

This section provides a comprehensive overview of the literature on SM, including its definitions and uses from 1994 to 2019, and offers recommendations for managers and researchers. The findings of this literature review demonstrate how the definitions of SM have evolved, concentrating increasingly less on individuals and their interactions and more on user-generated content and sharing. Social networks (SN), SM, and virtual communities (VC) are three groups into which definitions may be divided [66]. Just as a computer network is a collection of computers connected by a network of cables, an SN links individuals and groups. An SN is a collection of individuals (or organizations or other social entities) connected by a network of social interactions, such as friendships, coworking, or information sharing [67].

Social networking is becoming essential for finding and disseminating knowledge on diverse topics. Social media analytics is an emerging research area of immense interest to the information systems community. The goal of these analytics is to evaluate data from SM platforms; however, there are few comprehensive discussions or models available, and the field remains underdeveloped [67]. Social network analysts aim to comprehensively describe and analyze relationships, identify patterns, track information flow, and understand the impact of these networks on individuals and organizations [67]. Discovering, collecting, and preparing SM data poses many challenges to researchers, including the large volume and diver-

sity of data, in addition to the need for more advanced algorithms to process them [67] [68]. One proposed solution is to create advanced software structures, visual analyses, and innovative algorithms to help discover topics.

4.1. Rumors on Social Media

The examination of rumors on SM has become a major area of interest in SM analysis. Several researchers have examined the way in which unverified information moves in these digital spaces with the sole aim of understanding these dynamics and their effects. Such research has made significant contributions to the study of rumor propagation and has benefited our understanding of how false information moves in online spaces. Several studies have also contributed to describing the various dimensions in which rumors operate in the quest to explain the inter- or intraspecific relationships between human behaviors, network structures, and how information flows in these digital ecosystems.

The dynamics of rumor propagation have been investigated in multiple studies that have sought to reveal the strands of a rumor on the paths that people follow on SM networks. Many of these studies have used computational models, network analyses, and empirical studies to scrutinize how rumors start, gain, and lose momentum, spread across diversified user groups, and sometimes go viral. Most of these studies have employed network-based approaches and empirical analyses to study the patterns of rumor dissemination on X/Twitter. Network-based approaches, such as social network analysis, have identified influential users and network structures that provide avenues through which rumors can have longer lifespans [6].

Platforms such as Twitter/X and Reddit provide publicly available data, enabling researchers to more effectively study and track rumor spread patterns, user interactions, and content characteristics due to their open and accessible nature. In contrast, other social media platforms such as WhatsApp and Snapchat are important for rumor detection research due to their widespread use. Rumors on these platforms often spread through direct user interactions, making tracking difficult due to their closed nature, which limits the availability of public datasets, which are critical for rumor detection research. Researchers are therefore exploring innovative algorithms such as metadata analysis or leveraging trust mechanisms to address such challenges [69] [70].

Studies have demonstrated the varied characteristics of rumors about SM. For instance, rumors can go viral, which means they spread quickly and widely through SM platforms because users are all connected to each other [71]. Rumors can take advantage of the fact that the truth about a story is not known, making them more ambiguous and allowing individuals to form assumptions about the truth. Moreover, rumors are strengthened by individual beliefs, fears, or desires when shared through social circles. Rumors are also interesting because they change as they travel from friend to friend, or in our case, platform to platform; thus, they are harder to track and control [71]. To account for this semantic drift in rumors,

models require innovative techniques. For example, using temporal segmentations instead of random segmentations to train models on past data and test them on future data [72]. Dynamic embedding models, such as those based on word2vec or BERT, are also a good method for monitoring changes in word meanings and context over time [39]. Analyzing the path of rumor propagation by incorporating contextual implications of propagation paths provides a deeper understanding of rumor propagation, both structurally and semantically. Combining these two factors enhances the model's ability to adapt to the changing nature of rumors [61].

Furthermore, the spread of rumors is not solely reliant on their content; the structural properties of the underlying social networks also play a crucial role. Some methodologies have emphasized the need to debunk rumors in real time on social networks using the hybrid clustered shuffled frog-leaping algorithm with particle swarm optimization (SFLA-PSO), as well as taking advantage of trust relationships [58].

The present study, which examined friendship maintenance in social networks, highlighted the importance of trust in these digital relationships. The spread of rumors often threatens the dynamics of social networks; therefore, understanding how rumor propagation interferes with these dynamics is crucial [58] [73].

Both [58] and [73] made substantial contributions to the field by exploring the creation of a comprehensive trust mechanism. Their research highlights the significance of trust levels in online social network (OSN)-based social relationships, emphasizing this aspect over intimacy levels. Rumor detection in social networks can be a complex task because the state of information verification can change over time. Several recent studies have explored temporal classification to address the dynamic nature of rumors [74]. For example, the rumor detection method used in this study constructs a distributed spreading tree where the weight of each edge represents the temporal interval between connected posts. This method incorporates temporal information to model the spread of rumors over time, which improves the accuracy of rumor prediction and reduces the din [75].

As illustrated in **Figure 4**, at the rumor generation stage, a rumor frequently goes through numerous iterations of content within the social circle before making its way across social networks.

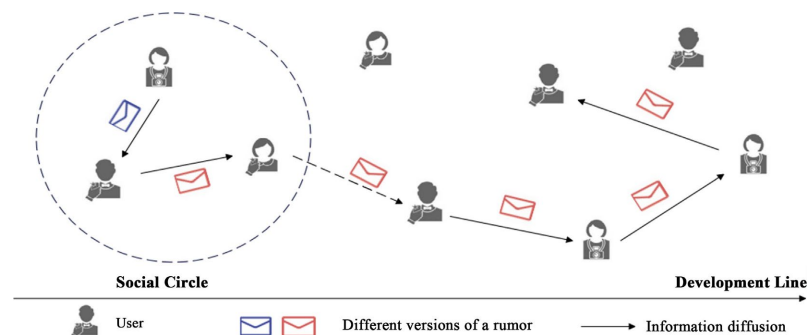


Figure 4. Information diffusion process during the rumor-generation phase.

4.2. Social Circle Mining (SCM)

Social circle mining is an analytical system that helps us understand the structure of social networks on digital platforms, such as online communities, and their functioning. SCM is used to extract, interpret, and analyze social circles from networks and employs various methodologies, including network science, ML, and data mining, to uncover communities and relationships within large networks [76]. The primary purpose of SCM is to identify and characterize clusters or groups of individuals with strong interconnections or shared interests within a wider social network [77] [78]. SCM uses methods including the traditional K-means clustering algorithm [79] and hierarchical clustering [80], as well as advanced algorithms such as community discovery algorithms based on module optimization or spectral clustering. Such methods uncover the structure of social networks, which leads to a deeper understanding of how they are organized [81] [82].

A practical example of this is the use of social circle mining [6], which extracts highly homogeneous social circles from user context, and combines them with social interaction to detect rumors. This method, called RDMSC, is processed and used in GCN encoders with an attention mechanism to learn and combine the two types of features for classification, as shown in Figure 5. The results demonstrate that the RDMSC method achieved superior rumor detection capabilities of up to 95%, outperforming state-of-the-art baselines on three real-world datasets.

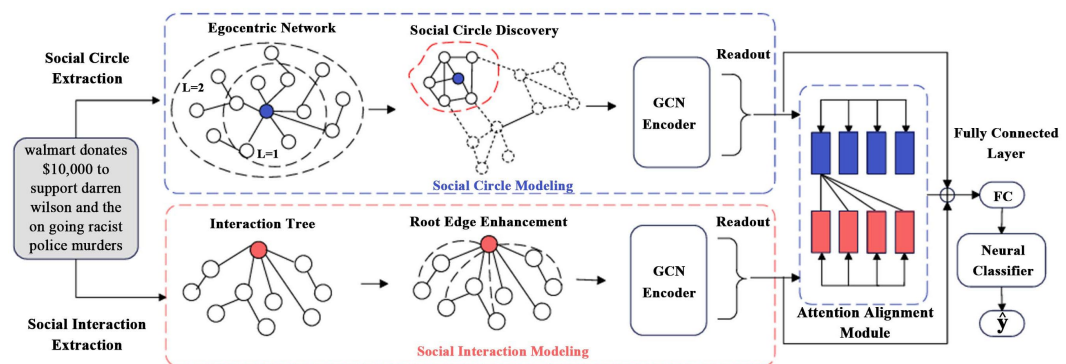


Figure 5. Overview of rumor detection model RDMSC.

Consider a practical example where SCM is integrated with a deep learning model for rumor detection. Data is collected from tweets related to a trending topic. The data is cleaned by removing retweets, filtering out non-English tweets, and normalizing text data. Hierarchical clustering is used to identify groups of users who frequently interact with each other. Features such as the number of interactions within each circle, sentiment of tweets, and frequency of keyword usage are extracted. These features are integrated into an LSTM (Long Short-Term Memory) network to capture temporal dependencies and patterns in the data. The LSTM network is trained to classify tweets as rumors or non-rumors based on the extracted features. The model's performance is evaluated using a validation da-

taset, and it is refined based on the results. Such examinations enable the study of various social media research problems, such as echo chambers or filter bubbles, how information is diffused, the important factors contributing to information becoming viral in each social circle, and the trajectory of information diffusion (such as rumor debunking) [83] [84]. Moreover, SCM also helps deliver targeted content because, through SCM, a platform can ascertain social circles based on user interaction and shared interests, such as movies, stock market information, or religious factors, enabling platforms to prompt collective behavior by targeting groups with those shared interests [82] [85]. Furthermore, SCM holds significant promise for rumor detection and mitigation. SCM is capable of understanding how rumors spread within a particular community by analyzing the structural properties of social circles, such as identifying influential nodes in a diffusion process and mapping diffusion paths [6] [84]. Furthermore, this capacity can facilitate the development of proactive strategies to intervene and counteract the spread of misinformation. It can also help mitigate its impact by targeting influential nodes and disrupting the pathways of rumor diffusion within these social circles.

5. Comparisons

We conducted a literature survey on rumor detection methods (Section 4), focusing on the use of social circle mining. We classified rumor detection methods into four categories: machine learning (ML), deep learning, hybrid, and network-based approaches. ML approaches are characterized by their speed and efficiency in processing large amounts of data. Their predictions may improve over time, but they may be complex to develop and maintain, as social network data are characterized by rapid development. Deep learning approaches can process large amounts of data with high accuracy and recognize complex patterns in the data. However, they require expensive computational resources. Hybrid approaches can provide the advantages of both categories by combining ML and deep learning methods. Most recent studies have adopted this approach to enhance efficiency in rumor detection and the ability of models to adapt to different patterns. However, hybrid approaches may be comparatively more complex and expensive. Finally, network-based approaches, including social network mining, provide broad and rapid access to large amounts of data and immediate updates. In addition to their ability to analyze interactions on social networks to help understand how rumors spread, such approaches may face challenges related to privacy, security, and obtaining irrelevant information.

Regarding rumor detection using transformer-based architectures, such as ChatGPT models, despite their powerful rumor detection capabilities, there are some ethical issues, including privacy concerns, as they require exceptionally large datasets for training, raising concerns about collecting this data. Moreover, the built-in bias in the data they were trained on led to discriminatory results in rumor detection. Furthermore, these transformers can be used maliciously to spread rumors and manipulate public opinion.

A technical comparison of these various approaches is provided in the following tables (**Tables 1-4**).

Table 1. Comparison of machine-learning approaches for rumor detection.

Study	Detection Method	Subcategory	Description/Approach	Dataset Used	Results/Evaluation
[30]	Extreme Gradient Boosting Method	Arabic Tweet Rumor Detection	Used XGBoost to detect rumors in Arabic tweets, employing a comprehensive feature set and achieving high accuracy.	Public Arabic Tweet Dataset	Achieved 97.18% accuracy.
[22]	Machine Learning Models for Virality Prediction	Fake News Virality Prediction	Predicted the virality of fake news at an early stage using textual features, distinguishing between true and fake news propagation.	Not specified	Discussed preventive strategies based on prediction accuracy.
[28]	Machine Learning Methods	Disaster-Related Rumor Refuters	Compared various ML techniques for identifying disaster-related anti-rumor spreaders, using sentiment and text similarity analysis for prediction.	Sina Weibo Users and Microblogs	Developed a robust XGBoost model for prediction.
[5]	Content and Novelty-Based Features	Rumor Prediction	Introduced a new task for predicting potential future rumors, combining content and novelty features to improve detection accuracy.	Not specified	Improved accuracy in rumor prediction.
[21]	Heterogeneous User Representation Model	Rumor Detection	Developed a model that uses heterogeneous user representation to differentiate between rumors and credible messages by analyzing information propagation patterns.	Not specified	Effective in distinguishing rumors based on propagation patterns.
[31]	XGBoost and NLP Techniques	Rumor Refuter Feature Analysis	Analyzed user features and sentiment using NLP and XGBoost to identify potential rumor refuters, showing differences across rumor categories.	Sina Weibo Users and Microblogs	Revealed significant feature differences among refuters.

Continued

[24]	Logistic Regression	Health Rumors	Investigated features distinguishing true from false health rumors using 453 health rumors and a logistic regression model.	Online reference of health rumors in China	Lengths of headlines/statements and presence of pictures negatively correlated with truthfulness; elements such as numbers and hyperlinks positively correlated. Dread rumors are more likely to be true.
[20]	Classification	Twitter Credibility	Assessed the credibility of tweets related to trending topics using features from posting behavior, text, and external citations.	Twitter postings related to trending topics	Precision and recall ranged from 70% to 80% for classifying credibility.
[25]	Regression & Classification	Popularity Prediction	Predicted future popularity of false rumors using post- and user-level information; introduced the BERT-Weibo-Rumor model.	19,256 false rumor cases from Weibo	KG-Fusion model outperformed baselines with an RMSE of 1.54 and Pearson's r of 0.636. Popular rumors used more conjunctions and punctuation.
Unsupervised-Based Approaches					
[33]	Rule-Based Detection	Twitter Rumors	Analyzed rumor spread after a disaster and developed a system to detect rumor candidates on Twitter.	Twitter data from disaster events	The proposed algorithm detected rumors with acceptable accuracy.
[35]	Rule-Based Detection	Political Rumors	Detected extreme political rumors on Twitter by identifying extreme users and using structural and timeline features.	Twitter data on political rumors	The rule-based method provided high precision and recall for detecting political rumors, with some rules achieving 100% precision.
[36]	Classification	Misinformation Detection	Detected misinformation on Twitter in real-time by comparing tweets from verified news channels and general users.	Twitter data and the Twitter Grapevine prototype	Large number of topics flagged as suspicious with high accuracy.

Continued

Self-supervised-Based Approaches					
[40]	Models BERT	Arabic Fake News Detection	Used mini-BERT with Arabic language embedding for fake news detection in Arabic, compared with ML classifiers.	Arabic fake news dataset	Mini-BERT-based classifiers outperformed ML classifiers, showing improved performance with increased training data.
[39]	Models BERT	Rumor Classification	Combined BERT-OPCNN for feature extraction with FIAC and Bi-LSTM for classification of rumor text.	Various rumor datasets	FIAC embedding with BERT-OPCNN outperformed existing techniques in rumor classification.
[38]	Contrastive Learning	Rumor Tracking	Proposed SimCLRT framework for tracking rumors using contrastive learning; included variants SimCLRT-CNN, SimCLRT-Linear, and SimCLRT-RNN.	PHEME and RumorEval datasets	SimCLRT models outperformed baselines; SimCLRT-CNN showed the best performance overall.

Table 2. Comparison of deep learning approaches for rumor detection.

Study	Detection Method	Subcategory	Description/Approach	Dataset Used	Results/Evaluation
[86]	Convolutional Neural Network (CNN)	Deep Learning (CNN)	Proposed CAMI, a CNN-based approach for misinformation identification and early detection.	Two large-scale datasets	CAMI effectively identified misinformation and achieved early detection.
[47]	Event Adversarial Neural Network (EANN)	Deep Learning (Hybrid)	Framework with multimodal feature extractor, fake news detector, and event discriminator.	Multimedia datasets (Weibo, Twitter)	EANN outperformed state-of-the-art methods in fake news detection and learned transferable features.
[3]	RNN-based Detection	Breaking News Rumors	Detected breaking news rumors using word embeddings and RNNs, addressing topic shift issues in rumor detection.	Real-life rumor dataset	The model outperformed state-of-the-art methods in precision, recall, and F1 score.

Continued

[44]	Tree-Structured RNN	Twitter Rumors	Detected rumors on Twitter using tree-structured recursive neural networks to handle sentiment expression and user bias.	Twitter datasets in English and Chinese	The method significantly outperformed baseline algorithms in detecting rumors.
[45]	Hybrid Deep Model	Fake News Detection	Combined text, user response, and source characteristics for fake news detection using the CSI model.	Real-world news data	The CSI model achieved higher accuracy than existing models and provided meaningful latent representations.
[57]	Gated Graph Neural Network (GNN)	Deep Learning (GNN)	Used a graph-based approach to model content and semantic dependencies for rumor detection.	Real-world datasets	Outperformed existing methods by addressing redundant information and improving detection accuracy.
[86]	Evidence-Providing Rumor Detection (EPRD)	Hybrid Approaches	Incorporated prior knowledge and current comments using GraphSAGE and attention mechanisms.	Two real-world Twitter datasets	Best performance in rumor detection and early detection compared with baselines.
[51]	Graph Neural Network (GNN)	Hybrid Approaches	Predicted viral rumors and vulnerable users using a unified GNN model with multitask learning.	Datasets with annotations	Showed significant improvements in rumor detection, virality prediction, and user vulnerability scoring.
[50]	Propagation Graph Neural Network (PGNN)	Deep Learning (GNN)	Constructed a propagation graph and used PGNN with attention mechanisms for rumor detection.	Real-world Twitter dataset	Achieved better performance than state-of-the-art methods for rumor detection and early detection tasks.
[49]	Bi-Directional Graph Convolutional Networks (Bi-GCN)	Deep Learning (GNN)	Used Bi-GCN to capture both propagation and dispersion characteristics of rumors.	Various benchmarks	Demonstrated superior performance over existing approaches in detecting rumors on social media.

Table 3. Comparison of hybrid approaches for rumor detection.

Study	Detection Method	Subcategory	Description/Approach	Dataset Used	Results/Evaluation
[40]	Mini-BERT-based Deep Learning Classifiers	Arabic Fake News Detection	Used mini-BERT for sentiment analysis of Arabic fake news and compared it with ML classifiers. Applied holdout validation schemes to evaluate performance.	Arabic Fake News Dataset	Mini-BERT-based classifiers outperformed ML classifiers.
[53]	Rumor Detection Framework	Social Media Rumors	Proposed a framework that visualizes topic structures in time series, extracts rumor candidates, and verifies them using additional media sources.	Not specified	Not specified
[55]	Hybrid CNN and RNN Models	Fake News Classification	Proposes a framework using a combination of CNN and RNN models to classify fake news on Twitter with an accuracy of 82%.	Twitter Posts Dataset	Achieved 82% accuracy.
[54]	Feature Extraction from Crowd Responses	Rumor Detection in Chinese	Focused on detecting rumors on Weibo by extracting features from retweets and comments, employing clustering analysis and a classifier based on observed feature distribution.	Weibo Dataset	New features improved classification effectiveness.
[59]	Unsupervised Learning Model	User Behavior-based Detection	Combined RNNs and autoencoders to detect rumors based on users' behaviors, achieving high accuracy and F1 measure.	Not specified	Accuracy of 92.49% and F1 measure of 89.16%.
[58]	Hybrid Clustered SFLA-PSO Algorithm	Rumor Refutation	Proposed a hybrid clustered shuffled frog-leaping algorithm-particle swarm optimization (HCSFLA-PSO) for timely and real-time rumor refutation, integrating new trust mechanisms and energy consumption models.	Not specified	Demonstrated effectiveness through numerical simulations.

Continued

[39]	Hybrid BERT-OPCNN & FIAC with Customized Bi-LSTM	Rumor Text Classification	Used a two-phase feature extraction approach combining BERT-OPCNN and FIAC with customized Bi-LSTM for rumor classification, showing improved performance over existing techniques.	Not specified	Outperformed existing techniques in classification.
[56]	Portable Graph Transformer-based Model	Multimodal Graph-Based Detection	Introduced PHAROS, a graph transformer-based model with multimodal homophily measures designed to handle heterophily and integrate label information.	Real and Synthetic Data	Demonstrated superiority, efficiency, and robustness.
[57]	Psychological Motivation-based Detection	Endogenous Psychological Motivation	Focused on the psychological motivations behind user behaviors in social networks for rumor detection, analyzing active and passive user responses.	Twitter16 Dataset	Improved accuracy by 2.1% over the baseline.
[38]	Contrastive Learning Framework	Rumor Tracking	Proposed SimCLRT, a contrastive learning framework for rumor tracking those addresses tweet coverage issues, with variants showing varying performance across datasets.	PHEME, RumorEval Datasets	Outperformed baselines, with SimCLRT-CNN performing best.

Table 4. Comparison of network approaches for rumor detection.

Study	Detection Method	Subcategory	Description/Approach	Dataset Used	Results/Evaluation
[60]	Evolutionary Game Model	Rumor Propagation Control	Analyzed rumor propagation and control in social networks using evolutionary game theory and simulation.	Real rumor dataset from Twitter	The model showed the impact of anti-rumor messages and social anxiety on rumor spread.
[61]	Optimal Control	Rumor Propagation	Examined rumor propagation with influential and ordinary nodes; proposed control strategies for homogeneous and heterogeneous networks.	Homogeneous and heterogeneous network datasets	Government policies and immunization strategies effectively reduced rumor spread.

Continued

[74]	Heterogeneous Modeling	Rumor Detection	Developed a new model for rumor detection based on heterogeneous user representation and information propagation patterns.	Social media data	The model effectively distinguished rumors from credible messages; rumors spread among specific user groups.
[6]	Homogeneity Mining	Social Circle Mining	Proposed a new algorithm for detecting rumors by exploring social circles with high homogeneity in user context.	Real-world social media datasets	The approach outperformed existing methods in detecting early-stage rumors.

6. Challenges and Future Trends

Detecting rumors on social networks faces several challenges, the most prominent of which is the large volume of data generated by these social networks, as analyzing this vast amount of data requires advanced algorithms and computational models. Furthermore, this data varies in its formats and includes text, images, video clips, and other forms, which further complicates the process of rumor analysis and detection.

Another challenge lies in understanding the context and meaning of the published texts, as colloquial dialects or cultural backgrounds are often relevant and can make it difficult to interpret such texts on systems, thereby leading to incorrect interpretations in detecting rumors. Additionally, spreading rumors and transmitting them from fake or fictitious accounts hinders the process of tracking rumors and ascertaining their correct sources.

The scarcity of labeled data required to train rumor detection models is another challenge. Obtaining, filtering, and classifying data is a tedious task that requires considerable effort and time but is necessary for training and developing such systems. Furthermore, rumors are characterized by their sophistication and ability to change rapidly over time, which requires continuous training and updating of models. Therefore, it is imperative to keep rumor detection models sophisticated and effective on an ongoing basis.

Social network mining approaches show tremendous potential for rumor detection, but they face several limitations that currently prevent their widespread use. These limitations include the availability of rich contextual data on user interactions and network structures, which makes collecting and processing them difficult at scale due to privacy and platform constraints. Second, due to the diverse types of rumors and platforms, detection models may face difficulties in generalizing, especially when dealing with new rumors. Third, rumors vary over time, so models based on static data may be less effective, resulting in the temporal drift of concepts. Furthermore, models designed for rumor detection may suffer from significant computational complexity, which may make their current use

challenging [87]. Furthermore, the problem of uneven data balance (verified rumors being less common than unverified ones) reduces the accuracy of models. Furthermore, these models require significant computing resources to be implemented on large-scale social media platforms [87]. Addressing these limitations requires significant improvements in data collection, the development and design of appropriate algorithms, and improved computing infrastructure.

Given these challenges, SCM for rumor detection is considered a distinctive and promising solution, given features such as network structure analysis, community detection, and the ability to explicate and understand behavioral patterns. Through this approach, it is possible to monitor the nodes that influence rumor spreading and evaluate the credibility of information published on the network by building trust levels as well as tracking the path of rumors. Despite the capabilities and features provided by SCM for rumor detection, studies in this field are limited. This indicates the need for more research and development in this field to enhance the detection of rumors on social networks.

7. Conclusion

Social media has gained popularity since its emergence and has become an integral part of daily life. These platforms have become hospitable environments for the generation and spread rumors. Rumors can cause significant harm to individuals, including reputational or financial loss and other damages. This survey provides a comprehensive view of the methods currently used to detect rumors on social networks, focusing on the use of SCM to detect rumors. Thus, this manuscript addresses numerous related concepts, such as the definition of rumors, methods of rumor detection, social networks, how rumors spread within them, and mining social circles to detect rumors. We classified the commonly employed methods into four categories: machine learning-based, deep learning-based, hybrid, and network-based approaches. Machine learning and deep learning models have proven capabilities in rumor detection. However, most of these methods ignore the impact of users' social circles on the spread of rumors. Such methods may also be challenging to handle and require continuous updates to vast amounts of data. Hybrid approaches are the most prominent and widespread methods used to take advantage of the features of all models, thereby enhancing their capabilities in detecting rumors. Studies have emphasized the importance of network-based approaches, as they consider the context and content of rumors and the spread mechanism, which is significant in detecting rumors that have been difficult for previous approaches to explain. Moreover, previous studies have emphasized the importance of SCM as an additional approach that enhances the rapid detection of rumors and provides a comprehensive view of the mechanisms of rumor spread. Furthermore, integrating SCM into traditional detection methods will allow the development of more flexible and efficient methods to address misleading information and disinformation in digital technology. However, more research is needed in this field.

Acknowledgements

The authors would like to thank the Deanship of Scientific Research, Qassim University, for funding the publication of this project.

Conflicts of Interest

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

References

- [1] Allcott, H. and Gentzkow, M. (2017) Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives*, **31**, 211-236. <https://doi.org/10.1257/jep.31.2.211>
- [2] Oyza, I. and M. Edwin, A. (2015) Effectiveness of Social Media Networks as a Strategic Tool for Organizational Marketing Management. *The Journal of Internet Banking and Commerce*, **1**. <https://doi.org/10.4172/1204-5357.s2-006>
- [3] Alkhodair, S.A., Ding, S.H.H., Fung, B.C.M. and Liu, J. (2020) Detecting Breaking News Rumors of Emerging Topics in Social Media. *Information Processing & Management*, **57**, Article ID: 102018. <https://doi.org/10.1016/j.ipm.2019.02.016>
- [4] Pathak, A.R., Mahajan, A., Singh, K., Patil, A. and Nair, A. (2020) Analysis of Techniques for Rumor Detection in Social Media. *Procedia Computer Science*, **167**, 2286-2296. <https://doi.org/10.1016/j.procs.2020.03.281>
- [5] Qin, Y., Dominik, W. and Tang, C. (2018) Predicting Future Rumours. *Chinese Journal of Electronics*, **27**, 514-520. <https://doi.org/10.1049/cje.2018.03.008>
- [6] Zheng, P., Huang, Z., Dou, Y. and Yan, Y. (2023) Rumor Detection on Social Media through Mining the Social Circles with High Homogeneity. *Information Sciences*, **642**, Article ID: 119083. <https://doi.org/10.1016/j.ins.2023.119083>
- [7] Mengist, W., Soromessa, T. and Legese, G. (2020) Method for Conducting Systematic Literature Review and Meta-Analysis for Environmental Science Research. *MethodsX*, **7**, Article ID: 100777. <https://doi.org/10.1016/j.mex.2019.100777>
- [8] Miller, D.E. (1992) "Snakes in the Greens" and Rumor in the Innercity. *The Social Science Journal*, **29**, 381-393. [https://doi.org/10.1016/0362-3319\(92\)90002-y](https://doi.org/10.1016/0362-3319(92)90002-y)
- [9] Rosnow, R.L. (1976) Rumor and Gossip: The Social Psychology of Hearsay. Elsevier.
- [10] Bock, P.G. (1964) A Theory of Rumor Transmission. *Public Opinion Quarterly*, **28**, 687-690. <https://doi.org/10.1086/267293>
- [11] Kapferer, J.N. (1990) Rumors: Uses, Interpretations, and Images. Routledge.
- [12] DiFonzo, N. and Bordia, P. (2007) Rumor, Gossip and Urban Legends. *Diogenes*, **54**, 19-35. <https://doi.org/10.1177/0392192107073433>
- [13] Zubiaga, M.L.A. and Procter, K.B.R.N. (2015) Towards Detecting Rumours in Social Media. *Proceedings of the AAAI Workshop*, Austin, 25-30 January 2015, 35-41.
- [14] Liu, X.F., Burton-Jones, A., Liu, D.B.J. and Xu, A. (2014) Rumors on Social MEDIA in Disasters: Extending Transmission to Retransmission. *Proceeding of the 19th Pacific Asia Conference on Information Systems*. <http://aisel.aisnet.org/pacis2014/49>
- [15] Ahsan, M., Kumari, M. and Sharma, T.P. (2019) Rumors Detection, Verification and Controlling Mechanisms in Online Social Networks: A Survey. *Online Social Networks and Media*, **14**, Article ID: 100050. <https://doi.org/10.1016/j.osnem.2019.100050>
- [16] Alpaydin, E. (2020) Introduction to Machine Learning. The MIT Press.

- [17] Jordan, M.I. and Mitchell, T.M. (2015) Machine Learning: Trends, Perspectives, and Prospects. *Science*, **349**, 255-260. <https://doi.org/10.1126/science.aaa8415>
- [18] Gongane, V.U., Munot, M.V. and Anuse, A. (2022) Machine Learning Approaches for Rumor Detection on Social Media Platforms: A Comprehensive Survey. In: Gupta, D., Sambyo, K., Prasad, M. and Agarwal, S., Eds., *Advanced Machine Intelligence and Signal Processing*, Springer, 649-663. https://doi.org/10.1007/978-981-19-0840-8_50
- [19] Doerr, B., Fouz, M. and Friedrich, T. (2012) Why Rumors Spread So Quickly in Social Networks. *Communications of the ACM*, **55**, 70-75. <https://doi.org/10.1145/2184319.2184338>
- [20] Castillo, C., Mendoza, M. and Poblete, B. (2011) Information Credibility on Twitter. Proceedings of the 20th international conference on World Wide Web, Hyderabad, 28 March-1 April 2011, 675-684. <https://doi.org/10.1145/1963405.1963500>
- [21] Liu, Y., Xu, S. and Tourassi, G. (2015) Detecting Rumors through Modeling Information Propagation Networks in a Social Media Environment. In: Agarwal, N., Xu, K. and Osgood, N., Eds., *Social Computing, Behavioral-Cultural Modeling, and Prediction*, Springer, 121-130. https://doi.org/10.1007/978-3-319-16268-3_13
- [22] Esteban-Bravo, M., Jiménez-Rubido, L.D.L.M. and Vidal-Sanz, J.M. (2024) Predicting the Virality of Fake News at the Early Stage of Dissemination. *Expert Systems with Applications*, **248**, Article ID: 123390. <https://doi.org/10.1016/j.eswa.2024.123390>
- [23] Bertsimas, D. and King, A. (2017) Logistic Regression: From Art to Science. *Statistical Science*, **32**, 367-384. <https://doi.org/10.1214/16-sts602>
- [24] Zhang, Z., Zhang, Z. and Li, H. (2015) Predictors of the Authenticity of Internet Health Rumours. *Health Information & Libraries Journal*, **32**, 195-205. <https://doi.org/10.1111/hir.12115>
- [25] Mu, Y., Niu, P., Bontcheva, K. and Aletras, N. (2024) Predicting and Analyzing the Popularity of False Rumors in Weibo. *Expert Systems with Applications*, **243**, Article ID: 122791. <https://doi.org/10.1016/j.eswa.2023.122791>
- [26] Noble, W.S. (2006) What Is a Support Vector Machine? *Nature Biotechnology*, **24**, 1565-1567. <https://doi.org/10.1038/nbt1206-1565>
- [27] Fernández-Delgado, E.C.M., Barro, D.A.S. and Fernández-Delgado, A. (2014) Do We Need Hundreds of Classifiers to Solve Real World Classification Problems? *Journal of Machine Learning Research*, **15**, 3133-3181.
- [28] Wang, S., Li, Z., Wang, Y. and Zhang, Q. (2019) Machine Learning Methods to Predict Social Media Disaster Rumor Refuters. *International Journal of Environmental Research and Public Health*, **16**, Article 1452. <https://doi.org/10.3390/ijerph16081452>
- [29] Chen, T. and Guestrin, C. (2016) XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, 13-17 August 2016, 785-794. <https://doi.org/10.1145/2939672.2939785>
- [30] Gumaei, A., Al-Rakhami, M.S., Hassan, M.M., De Albuquerque, V.H.C. and Camacho, D. (2022) An Effective Approach for Rumor Detection of Arabic Tweets Using Extreme Gradient Boosting Method. *ACM Transactions on Asian and Low-Resource Language Information Processing*, **21**, 1-16. <https://doi.org/10.1145/3461697>
- [31] Li, Z., Zhang, Q., Wang, Y. and Wang, S. (2020) Social Media Rumor Refuter Feature Analysis and Crowd Identification Based on XGBoost and NLP. *Applied Sciences*, **10**, Article 4711. <https://doi.org/10.3390/app10144711>

- [32] Dey, A. (2016) Machine Learning Algorithms: A Review. *International Journal of Computer Science and Information Technology*, **7**, 1174-1179.
- [33] Takahashi, T. and Igata, N. (2012) Rumor Detection on Twitter. *The 6th International Conference on Soft Computing and Intelligent Systems, and the 13th International Symposium on Advanced Intelligence Systems*, Kobe, 20-24 November 2012, 452-457. <https://doi.org/10.1109/scis-isis.2012.6505254>
- [34] Timbers, T., Campbell, T. and Lee, M. (2022) Data Science. Chapman and Hall/CRC. <https://doi.org/10.1201/9781003080978>
- [35] Chang, C., Zhang, Y., Szabo, C. and Sheng, Q.Z. (2016) Extreme User and Political Rumor Detection on Twitter. In: Li, J., Li, X., Wang, S., Li, J. and Sheng, Q., Eds., *Advanced Data Mining and Applications*, Springer, 751-763. https://doi.org/10.1007/978-3-319-49586-6_54
- [36] Jain, S., Sharma, V. and Kaushal, R. (2016) Towards Automated Real-Time Detection of Misinformation on Twitter. 2016 *International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Jaipur, 21-24 September 2016, 2015-2020. <https://doi.org/10.1109/icacci.2016.7732347>
- [37] Gao, Y., Wang, X., He, X., Feng, H. and Zhang, Y. (2022) Rumor Detection with Self-Supervised Learning on Texts and Social Graph. *Frontiers of Computer Science*, **17**, Article No. 174611. <https://doi.org/10.1007/s11704-022-1531-9>
- [38] Zeng, H. and Cui, X. (2022) SimCLRT: A Simple Framework for Contrastive Learning of Rumor Tracking. *Engineering Applications of Artificial Intelligence*, **110**, Article ID: 104757. <https://doi.org/10.1016/j.engappai.2022.104757>
- [39] Nithya, K., Krishnamoorthi, M., Easwaramoorthy, S.V., C R, D., Yoo, S. and Cho, J. (2024) Hybrid Approach of Deep Feature Extraction Using BERT-OPCNN & FIAC with Customized Bi-LSTM for Rumor Text Classification. *Alexandria Engineering Journal*, **90**, 65-75. <https://doi.org/10.1016/j.aej.2024.01.056>
- [40] Alawadh, H.M., Alabrah, A., Meraj, T. and Rauf, H.T. (2023) Attention-Enriched Mini-Bert Fake News Analyzer Using the Arabic Language. *Future Internet*, **15**, Article 44. <https://doi.org/10.3390/fi15020044>
- [41] Pathak, A.R., Pandey, M. and Rautaray, S. (2018) Application of Deep Learning for Object Detection. *Procedia Computer Science*, **132**, 1706-1717. <https://doi.org/10.1016/j.procs.2018.05.144>
- [42] Tan, L., Wang, G., Jia, F. and Lian, X. (2022) Research Status of Deep Learning Methods for Rumor Detection. *Multimedia Tools and Applications*, **82**, 2941-2982. <https://doi.org/10.1007/s11042-022-12800-8>
- [43] Cao, J., Guo, J., Li, X., Jin, Z., Guo, H. and Li, J. (2018) Automatic Rumor Detection on Microblogs: A Survey. arXiv: 1807.03505.
- [44] Ma, J., Gao, W. and Wong, K. (2018) Rumor Detection on Twitter with Tree-Structured Recursive Neural Networks. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Melbourne, 15-20 July 2018. 1980-1989. <https://doi.org/10.18653/v1/p18-1184>
- [45] Ruchansky, N., Seo, S. and Liu, Y. (2017) CSI: A Hybrid Deep Model for Fake News Detection. *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, Singapore, 6-10 November 2017, 797-806. <https://doi.org/10.1145/3132847.3132877>
- [46] Yu, F., Liu, Q., Wu, S., Wang, L. and Tan, T. (2017) A Convolutional Approach for Misinformation Identification. *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, Melbourne, 19-25 August 2017, 3901-3907.

- <https://doi.org/10.24963/ijcai.2017/545>
- [47] Wang, Y., Ma, F., Jin, Z., Yuan, Y., Xun, G., Jha, K., *et al.* (2018) EANN: Event Adversarial Neural Networks for Multi-Modal Fake News Detection. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, London, 19-23 August 2018, 849-857. <https://doi.org/10.1145/3219819.3219903>
- [48] Chen, J., Zhang, W., Ma, H. and Yang, S. (2023) Rumor Detection in Social Media Based on Multi-Hop Graphs and Differential Time Series. *Mathematics*, **11**, Article 3461. <https://doi.org/10.3390/math11163461>
- [49] Bian, T., Xiao, X., Xu, T., Zhao, P., Huang, W., Rong, Y., *et al.* (2020) Rumor Detection on Social Media with Bi-Directional Graph Convolutional Networks. *Proceedings of the AAAI Conference on Artificial Intelligence*, **34**, 549-556. <https://doi.org/10.1609/aaai.v34i01.5393>
- [50] Wu, Z., Pi, D., Chen, J., Xie, M. and Cao, J. (2020) Rumor Detection Based on Propagation Graph Neural Network with Attention Mechanism. *Expert Systems with Applications*, **158**, Article ID: 113595. <https://doi.org/10.1016/j.eswa.2020.113595>
- [51] Zhang, X. and Gao, W. (2024) Predicting Viral Rumors and Vulnerable Users with Graph-Based Neural Multi-Task Learning for Infodemic Surveillance. *Information Processing & Management*, **61**, Article ID: 103520. <https://doi.org/10.1016/j.ipm.2023.103520>
- [52] Yan, Y., Wang, Y. and Zheng, P. (2023) A Graph-Based Pivotal Semantic Mining Framework for Rumor Detection. *Engineering Applications of Artificial Intelligence*, **118**, Article ID: 105613. <https://doi.org/10.1016/j.engappai.2022.105613>
- [53] Hashimoto, T., Kuboyama, T. and Shirota, Y. (2011) Rumor Analysis Framework in Social Media. *TENCON2011—2011 IEEE Region 10 Conference*, Bali, 21-24 November 2011, 133-137. <https://doi.org/10.1109/tencon.2011.6129078>
- [54] Cai, G., Wu, H. and Lv, R. (2014) Rumors Detection in Chinese via Crowd Responses. *2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014)*, Beijing, 17-20 August 2014, 912-917. <https://doi.org/10.1109/asonam.2014.6921694>
- [55] Ajao, O., Bhowmik, D. and Zargari, S. (2018) Fake News Identification on Twitter with Hybrid CNN and RNN Models. *Proceedings of the 9th International Conference on Social Media and Society*, Copenhagen, 18-20 July 2018, 226-230. <https://doi.org/10.1145/3217804.3217917>
- [56] Nguyen, T.T., Ren, Z., Nguyen, T.T., Jo, J., Nguyen, Q.V.H. and Yin, H. (2024) Portable Graph-Based Rumour Detection against Multi-Modal Heterophily. *Knowledge-Based Systems*, **284**, Article ID: 111310. <https://doi.org/10.1016/j.knosys.2023.111310>
- [57] Yan, Y., Wang, Y. and Zheng, P. (2023) Rumor Detection on Social Networks Focusing on Endogenous Psychological Motivation. *Neurocomputing*, **552**, Article ID: 126548. <https://doi.org/10.1016/j.neucom.2023.126548>
- [58] Hu, X., Xiong, X., Wu, Y., Shi, M., Wei, P. and Ma, C. (2023) A Hybrid Clustered SFLA-PSO Algorithm for Optimizing the Timely and Real-Time Rumor Refutations in Online Social Networks. *Expert Systems with Applications*, **212**, Article ID: 118638. <https://doi.org/10.1016/j.eswa.2022.118638>
- [59] Chen, W., Zhang, Y., Yeo, C.K., Lau, C.T. and Lee, B.S. (2018) Unsupervised Rumor Detection Based on Users' Behaviors Using Neural Networks. *Pattern Recognition Letters*, **105**, 226-233. <https://doi.org/10.1016/j.patrec.2017.10.014>
- [60] Askarizadeh, M., Tork Ladani, B. and Manshaei, M.H. (2019) An Evolutionary Game Model for Analysis of Rumor Propagation and Control in Social Networks. *Physica A: Statistical Mechanics and its Applications*, **523**, 21-39.

- <https://doi.org/10.1016/j.physa.2019.01.147>
- [61] Myilsamy, K., Kumar, M.S. and Kumar, A.S. (2024) Optimal Control of a Rumor Propagation Model in Online Social Network by Considering Influential Nodes. *Results in Control and Optimization*, **14**, Article ID: 100339. <https://doi.org/10.1016/j.rico.2023.100339>
- [62] Khawaja, F.R., Zhang, Z., Memon, Y. and Ullah, A. (2024) Exploring Community Detection Methods and Their Diverse Applications in Complex Networks: A Comprehensive Review. *Social Network Analysis and Mining*, **14**, Article No. 115. <https://doi.org/10.1007/s13278-024-01274-1>
- [63] Anggrainingsih, R., Hassan, G.M. and Datta, A. (2024) Transformer-Based Models for Combating Rumours on Microblogging Platforms: A Review. *Artificial Intelligence Review*, **57**, Article No. 212. <https://doi.org/10.1007/s10462-024-10837-9>
- [64] Chang, Q., Li, X. and Duan, Z. (2024) A Novel Approach for Rumor Detection in Social Platforms: Memory-Augmented Transformer with Graph Convolutional Networks. *Knowledge-Based Systems*, **292**, Article ID: 111625. <https://doi.org/10.1016/j.knosys.2024.111625>
- [65] Lv, J., Wang, X. and Shao, C. (2022) TMIF: Transformer-Based Multi-Modal Interactive Fusion for Automatic Rumor Detection. *Multimedia Systems*, **29**, 2979-2989. <https://doi.org/10.1007/s00530-022-00916-8>
- [66] Aichner, T., Grünfelder, M., Maurer, O. and Jegeni, D. (2021) Twenty-five Years of Social Media: A Review of Social Media Applications and Definitions from 1994 to 2019. *Cyberpsychology, Behavior, and Social Networking*, **24**, 215-222. <https://doi.org/10.1089/cyber.2020.0134>
- [67] Garton, L., Haythornthwaite, C. and Wellman, B. (2006) Studying Online Social Networks. *Journal of Computer-Mediated Communication*, **3**, JCMC313. <https://doi.org/10.1111/j.1083-6101.1997.tb00062.x>
- [68] Stieglitz, S., Mirbabaie, M., Ross, B. and Neuberger, C. (2018) Social Media Analytics—Challenges in Topic Discovery, Data Collection, and Data Preparation. *International Journal of Information Management*, **39**, 156-168. <https://doi.org/10.1016/j.ijinfomgt.2017.12.002>
- [69] Albeshar, A.S. and Alhussain, T. (2021) Evaluating and Comparing the Usability of Privacy in WhatsApp, Twitter, and Snapchat. *International Journal of Advanced Computer Science and Applications*, **12**, 251-259. <https://doi.org/10.14569/ijacsa.2021.0120829>
- [70] Li, Q., Zhang, Q., Si, L. and Liu, Y. (2019) Rumor Detection on Social Media: Datasets, Methods and Opportunities. *Proceedings of the Second Workshop on Natural Language Processing for Internet Freedom: Censorship, Disinformation, and Propaganda*, Hong Kong, November 2019, 66-75. <https://doi.org/10.18653/v1/d19-5008>
- [71] Buchanan, T. (2020) Why Do People Spread False Information Online? The Effects of Message and Viewer Characteristics on Self-Reported Likelihood of Sharing Social Media Disinformation. *PLOS ONE*, **15**, e0239666. <https://doi.org/10.1371/journal.pone.0239666>
- [72] Mu, Y., Bontcheva, K. and Aletras, N. (2023) It's about Time: Rethinking Evaluation on Rumor Detection Benchmarks Using Chronological Splits. *Findings of the Association for Computational Linguistics: EACL 2023*, Dubrovnik, 2-6 May 2023, 736-743. <https://doi.org/10.18653/v1/2023.findings-eacl.55>
- [73] Liu, F. and Li, M. (2018) A Game Theory-Based Network Rumor Spreading Model: Based on Game Experiments. *International Journal of Machine Learning and Cybernetics*, **10**, 1449-1457. <https://doi.org/10.1007/s13042-018-0826-5>

- [74] Liu, Y. and Xu, S. (2016) Detecting Rumors through Modeling Information Propagation Networks in a Social Media Environment. *IEEE Transactions on Computational Social Systems*, **3**, 46-62. <https://doi.org/10.1109/tcss.2016.2612980>
- [75] Peng, X., Wu, J., Liu, R. and Xu, K. (2024) Rumor Detection on Social Media with Temporal Propagation Structure Optimization. arXiv: 2412.08316.
- [76] Leão, J.C., Brandão, M.A., Vaz de Melo, P.O.S. and Laender, A.H.F. (2018) Who Is Really in My Social Circle? Mining Social Relationships to Improve Detection of Real Communities. *Journal of Internet Services and Applications*, **9**, Article No. 20. <https://doi.org/10.1186/s13174-018-0091-6>
- [77] Wang, M., Zuo, W. and Wang, Y. (2016) An Improved Density Peaks-Based Clustering Method for Social Circle Discovery in Social Networks. *Neurocomputing*, **179**, 219-227. <https://doi.org/10.1016/j.neucom.2015.11.091>
- [78] Verbeke, W. and Wuyts, S. (2006) Moving in Social Circles—Social Circle Membership and Performance Implications. *Journal of Organizational Behavior*, **28**, 357-379. <https://doi.org/10.1002/job.423>
- [79] Iswed, I.M., F., Y. and S., A. (2019) Boosted Constrained K-Means Algorithm for Social Networks Circles Analysis. *International Journal of Advanced Computer Science and Applications*, **10**, 419-423. <https://doi.org/10.14569/ijacsa.2019.0100758>
- [80] Wang, C., Hao, C. and Guan, X. (2020) Hierarchical and Overlapping Social Circle Identification in Ego Networks Based on Link Clustering. *Neurocomputing*, **381**, 322-335. <https://doi.org/10.1016/j.neucom.2019.11.080>
- [81] Zhang, X., Ma, Z., Zhang, Z., Sun, Q. and Yan, J. (2018) A Review of Community Detection Algorithms Based on Modularity Optimization. *Journal of Physics: Conference Series*, **1069**, Article ID: 012123. <https://doi.org/10.1088/1742-6596/1069/1/012123>
- [82] Ren, S., Zhang, S. and Wu, T. (2020) An Improved Spectral Clustering Community Detection Algorithm Based on Probability Matrix. *Discrete Dynamics in Nature and Society*, **2020**, Article ID: 4540302. <https://doi.org/10.1155/2020/4540302>
- [83] Horne, B.D., Nørregaard, J. and Adalı, S. (2019) Different Spirals of Sameness: A Study of Content Sharing in Mainstream and Alternative Media. *Proceedings of the International AAAI Conference on Web and Social Media*, **13**, 257-266. <https://doi.org/10.1609/icwsm.v13i01.3227>
- [84] Wang, L. and Guo, Y. (2019) An Evolution Model of Rumor Spreading Based on WeChat Social Circle. *Journal of Information Processing Systems*, **15**, 1179-1191.
- [85] Palsetia, D., Patwary, M.M.A., Agrawal, A. and Choudhary, A. (2014) Excavating Social Circles via User Interests. *Social Network Analysis and Mining*, **4**, Article No. 170. <https://doi.org/10.1007/s13278-014-0170-z>
- [86] Li, J., Li, R., Ni, S. and Kao, H. (2024) EPRD: Exploiting Prior Knowledge for Evidence-Providing Automatic Rumor Detection. *Neurocomputing*, **563**, Article ID: 126935. <https://doi.org/10.1016/j.neucom.2023.126935>
- [87] Sattarov, O. and Choi, J. (2024) Detection of Rumors and Their Sources in Social Networks: A Comprehensive Survey. *IEEE Transactions on Big Data*. <https://doi.org/10.1109/tbdata.2024.3522801>