IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 17, 2017, accepted February 13, 2018, date of publication February 27, 2018, date of current version April 4, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2810062

Social Media Recommender Systems: Review and Open Research Issues

ANITHA ANANDHAN[1], LIYANA SHUIB[1], MAIZATUL AKMAR ISMAIL[1], AND GHULAM MUJTABA[1,2]

[1]Department of Information Systems, Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur 50603, Malaysia
[2]Department of Computer Science, Sukkur IBA University, Sukkur 65200, Pakistan

Corresponding authors: Anitha Anandhan (anitha.anandhan@gmail.com), Liyana Shuib (liyanashuib@um.edu.my), and Maizatul Akmar Ismail (maizatul@um.edu.my)

This work was supported by the University of Malaya's Research Grant Programme—SBS (Equitable Society)—under Grant RP059E-17SBS.

ABSTRACT In recent years, different types of review systems have been developed with the recommender system (RS). RSs are developed based on user textual reviews, ratings, and comparative opinions. RSs for social media resources, such as blogs, forums, social network websites, social bookmarking websites, video portals, and chat portals help users to collaborate effectively. Social media resources are used in the RS for recommending contents, articles, news, e-commerce products, and users. Although research on social media in RSs has increased annually, comprehensive literature review and classification of these RS studies are limited and must, therefore, be improved. This paper aims to provide a comprehensive review of the social media RS on research articles published from 2011 to 2015 by exploiting a methodological decision analysis in six aspects, including recommendation approaches, research domains, and data sets used in each domain, data mining techniques, recommendation type, and the use of performance measures. A total of 61 articles are reviewed among the initial 434 articles on RS research published in Web of Science and Scopus between 2011 and 2015. To accomplish the aim of this paper, a comprehensive review and analysis was performed on extracted articles to explore various recommendation approaches which are used in the RS. In addition, various social media domains are identified, where RSs have been employed. In each identified domain, publicly available data sets are also reported. Furthermore, various data mining techniques, recommendation types, and performance measures are also analyzed and reviewed in technical aspects. Finally, potential open research directions are also presented for future researchers intended to work in social media RS domain.

INDEX TERMS Recommender system, social network, social media, blog, forum, data mining technique, classification.

I. INTRODUCTION

In the era of fast-growing social media revolution, various computer-mediated technologies are used for online communication. These advancements are used as channels for community-based input, user interaction, creation, sharing information, real-life event dissemination, and private and public message exchanges via virtual communities and networks. Social media technologies are available in different forms, such as blogs, business networks (e.g., e-commerce), enterprise social networks, common question answering (CQA) forums, microblogs, photo-sharing portals, product review portals, social bookmarking, social gaming, social networks, video sharing, and virtual worlds. These advancements are used extensively to communicate among different groups of people in the world [1].

Social media website is defined as "a website that facilitates meeting people, finding like minds, communicating and sharing content, and building community"; this kind of website allows or encourages various types of activities, such as commercial, social, or a combination of the two [15]. Social media categories include digital library, e-commerce, entertainment, forum, geolocation, social bookmark, social review, social game, and social network. Social network is the subcategory of social media [43], which is the social structure of people who are joined by common interest (Figure 1). Social media are social channels of

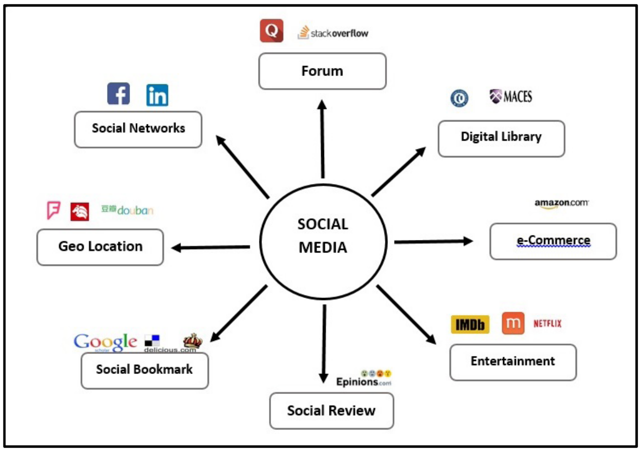


FIGURE 1. Categories of social media.

communication using web-based technologies, desktop computers, and mobile technologies. These technologies create highly interactive platforms through which individuals, communities, and organizations can share information, discuss, rate, comment, and modify user-generated and online contents. These advancements enable communication among businesses, organizations, communities, and individuals. Social media technologies change the way individuals and large organizations communicate, and they are increasingly being developed [2].

The following are examples of social media categories commonly used to connect people.

A. DIGITAL LIBRARY

A digital library is a collection of documents in organized electronic form and available in the Internet. A user may access magazine articles, books, papers, images, sound files, and videos, depending on specific library. Many organizations have established a digital library as an effective approach for academic institutions, museums, and certain city libraries.

B. FORUMS

In forums, such as Stack Overflow and Yahoo! Answers, users interact by posting (asking questions and soliciting opinions) and communicating with other users using a forum. In the context of blog comments, which are comments made by the audience at the end of a blog post, the interactions focus on the topic of the post in question.

C. E-COMMERCE

Commercial online transaction sites, such as Amazon, support social interaction and user contributions to assist online buying and selling of products and services.

D. ENTERTAINMENT

Entertainment sites support interactive functionality and content, including live video streaming, movies, video chat communications, and music and videos streaming. Examples of this type are MovieLens, MUADDIB, IMDB, Smart Television, Netflix, Flixster, and Jester dataset (jokes).

E. BLOG/MICROBLOG

In Twitter, an example of a blog/microblog, users share updates with the people subscribed to them. These posts are short and often limited with a specific word count.

F. SOCIAL NETWORK

Facebook (social) and LinkedIn (professional) allow users to make virtual social connections with others by providing the option to set up one's own profile in social network sites. Users access different features of the network to connect and share with others and interact in other ways.

G. GEOLOCATION

Geolocation application, such as Foursquare, enable users to share their current location with other members of their social network and earn virtual badges for checking into sites.

H. SOCIAL GAMES

Online games allow or require social interaction between players, as opposed to playing games in solitude. Examples of social games are pogo, card games, and board games.

I. SOCIAL BOOKMARKING

Delicious, CiteULike with Google Scholar, and BibSonomy are sites where users can tag or bookmark different webpages and sites they like for future use. Such media provide different features to organize and manage multiple links and share them publicly.

J. SOCIAL REVIEW

This category, such as Epinions, is a general consumer review site where customers/users can read new and old reviews about a variety of items to help them decide on a purchase.

K. OTHERS (PUBLISH, SHARE, AND DISCUSS)

This category allows users to interact with other users by social media profiles, messages, and comments. Moreover, this type permits users to share their content (images and videos). YouTube, Flickr, Skype, Gtalk, and Wikipedia are some examples under this category.

Nowadays, a large amount of information (including data, images, videos, contents, and documents) is shared on social media. This considerable sharing of information introduces the problem of information overload to users. Therefore, several social media websites utilize recommender systems (RSs) to overcome this issue and suggest useful information to targeted users. RSs have been extensively explored in the mid-1990s [3]. RSs are used for recommending books, movies, music, products, and television programs. RSs analyze suggestions from users [2] in the form of reviews and ratings [70]. In the academic context, the RSs of scientific and online libraries support users by allowing them to move beyond catalog searches using efficient and accurate recommendation techniques within the systems, thereby obtaining relevant and dependable recommendations. RSs can reduce the time and cost spent by users and improve the process, quality, and decision-making strategy for providers [3].

Social media RSs generate a useful recommendation of articles and items that helps users collaborate with other users. Social networks (Facebook, Twitter, and blogs) contain a large quantity of information available as online documents and archives. Social networks aim to allow connection among friends and contain valuable information about user preferences. However, contents shared on these networks are noisy and heterogeneous, and they must be processed for information extraction [7]. Users discuss their specialization and highlight their opinions online. The number of online items is rapidly growing. Thus, determining items in a particular specialization becomes difficult for experts. RS is an essential solution to this problem and used in social media sites to identify the neighbors of target users based on user profile. These systems suggest the target user items liked or posted by neighboring users [6].

Identifying the suitable item or information in the internet becomes challenging due to information overload. RS is an information filtering system that deals with this problem by filtering primary information according to user preferences, interests, liked items, and ratings on the preferred item. These factors predict whether the user will prefer the item or information according to his or her profile. Social media grow at a high rate as a result of the popularity and large number of users in the network. RSs are continuously being explored and developed because of information overload and popularity of social media. Nevertheless, information overload is significantly challenging because it creates new research issues in other social media categories, such as large dataset in forums and digital libraries of books and documents. Geolocation includes tourism and travel and involves areas of entertainment datasets in social media [11]. RSs enhance the ratings and revenues for effectiveness of items or products. These research issues receive increasing attention in social media RS. In addition, the performance and data mining techniques of RS are unexplored.

In general, the four different approaches used in developing the RSs include content-based (CB) filtering, collaborative filtering (CF), hybrid-based (HB) filtering, and knowledge-based (KB) filtering. In CB filtering, recommendations are made using features and preferences. Content similarity is used rather than that of users [6]. CF defines similar users as active users. Opinions and previously stated interests of like-minded users are used in this system. A combination of CB and CF methods is known as HB filtering, which is the most remarkable approach that eliminates the drawbacks of CB and CF [5]. KB filtering is based on explicit knowledge of an item assortment, user preferences, and recommendation criteria.

In social media RSs, several data mining techniques, such as k-means, k-nearest neighbor (*k*NN) are utilized to reveal the hidden patterns or interests of users from a large amount of data [4]. Such techniques are used in social media RSs to obtain insights from data and provide useful recommendations. In general, two approaches are used in RSs in social media to provide useful recommendations based on ratings or rankings, namely, classical or two dimensional and contextual or multidimensional. Generally, classical approaches include recommendation types, such as item–user (relevance between pair of item and user), user–item (relevance between pair of user and item), user–tag (relevance between a pair of user and tag), and user–user (relevance between a pair of user and user). Contextual or multidimensional approaches, such as user–tag–items or user–item–tag, use more than two dimensions to provide useful recommendations. Various performance metrics are used to measure the performance of RS techniques, such as precision, recall, and accuracy. Precision measures the efficiency of relevant items selected. Recall is the number of relevant items selected and the total number of items in whole set. Accuracy measures how well the user ratings can be reproduced.

In the past few years, many reviews and surveys were published on RS in social media [48], [58]. These studies focused on the traditional approaches, such as CB, CF, and HB, and provided an overview of existing technologies for establishing personalized RSs and recommendation approaches with techniques [51], [61], [84]. However, these published reviews are only limited to e-business application domains, forums, bookmarking websites, and social networks [11], [38], [46], [56], [60]. Moreover, these published reviews mostly summarized the existing work in the field of social media RSs from the perspective of recommendation approaches and types [49], [67]. The present review is different from existing reviews by focusing on RS according to six aspects. This study aims to review and analyze six aspects of social media RSs, namely, (1) recommendation approaches, (2) social media subdomains, (3) datasets, (4) data mining techniques, (5) recommendation types, and (6) performance metrics. The specific contributions of this review are listed below:

1. This review provides a comprehensive overview of social media, including the subdomains of social media, and the role of RSs in social media.
2. This review will also provide a comprehensive review of recommendation approaches used to design the RSs.
3. Various social media subdomains are reviewed. In these subdomains, RSs are designed to provide useful recommendations to users. In each subdomain, a variety of datasets used and freely available for future researchers are identified and reviewed.
4. Several data mining techniques used to discover the users' interests or preferences about the items or users are reviewed thoroughly.
5. Various recommendation types, such as classical and contextual, used in selected studies are rigorously analyzed and reviewed.
6. The widely used performance metrics used to evaluate the performance of RSs are also identified and reviewed.

7. Finally, social media RSs are still in the development stage, and many open issues need further investigation. Thus, this review presents possible research directions to improve the performance of social media recommendation in different social media domains.

In this review, 61 primary studies were selected after a rigorous selection process from two large academic databases, namely, Web of Science (WOS) and Scopus. This review will answer the following research questions.

(1) What recommendation approaches are used in social media RS?
(2) What data mining techniques are used in social media RS?
(3) What datasets are used in social media RS?
(4) What recommendation types are preferred in social media RS?
(5) How is social media RS evaluated?

The remaining part of this paper is organized as follows. Section 2 presents the research methodology. Section 3 analyzes and discusses the bibliometric analysis of social media RS. Section 4 presents the review and analysis of social media recommendation approaches, domains, datasets, data mining techniques, recommendation types, and performance metrics. Section 5 presents the discussion on these six aspects. Potential future research directions in the field of social media RSs are presented in Section 6. Finally, Section 7 concludes this review.

II. RESEARCH METHODOLOGY

This study aims to understand the trend of social media RSs by examining published articles within the year limit. The steps of the research methodology are summarized as follows.

- Classification of research papers on social media RSs.
- Analysis of research papers on social media RSs and presentation of results based on classification.
- Bibliometric analysis of results based on classifications such as year-wise, country-wise, and journal-wise distributions.
- Analysis of importance of social media RSs and improvements.
- Discussion, conclusion and limitations of the study.

Web of Science (WOS) and Scopus databases are used in this study for selecting the articles. WOS and Scopus are authentic databases and cover many other science-related databases. The two databases can help analyze the scientific field articles and ease the analysis. The use of these databases allows searching and sorting the results. The results are extracted by expected parameters, such as first author, citation, and institution. Regarding the impact factor and h-index of the two databases, different results are obtained from existing studies. The perceptions of authors and researchers on both databases can be investigated to determine which of the two is best [74], [76]. WOS and Scopus have salient features of provenance and coverage, searching and analysis of results, citation tracking and analysis, impact factor, indexing (h-index), researcher profiles, and tools. Prior studies that compare these databases will be reviewed in the analysis part.

Although, the search query in Google Scholar retrieves lot of academic literature, however, all these publications may not justify or maintain the quality standards. There are various reasons for selecting papers only from Web of Science and Scopus databases and not selecting from Google scholar which are stated as follows.

1. **Quality:** There are variety of journals which are low quality and are indexed in Google Scholar. Conversely high quality journals only are indexed in WOS, and Scopus. Moreover, both these databases are large which index almost all well-reputed publishers.
   Hence, if we include the research papers from Google Scholar, the quality of our review paper may compromise.
2. **Peer review**: In WOS and Scopus databases all the indexed journals are peer-reviewed quality journals [73]. However, there are various journals which are indexed in Google scholar and they claim that these journals are peer-reviewed journals, however, in reality they just take one week to review, accept, and publish the submitted papers. Therefore, their peer review process in questionable.
3. **Accessibility**: Though, Google Scholar is free access search engine [74], it indexes the publications record only if the publisher willing to provide along with the abstract which is freely accessible otherwise it is subscription based. However, if one has access to WOS or Scopus databases, he/she can download the full text including all cited references.

The distribution of research papers on social media RSs is verified by their year of publication and classification based on recommendation approach, data mining technique, domain, datasets, recommendation type, and performance metric. However, confining each paper discussing social media RSs to a specific discipline is difficult because research papers on social media RSs are scattered across diverse domains, such as geolocation, forum, entertainment, e-commerce, social review, social network and social bookmark. As a result, the increasing number of research papers on RSs must be compiled. The search process of research papers on social media RSs is performed on 61 journals totally, the top 47 journals of WOS and 14 journals of Scopus.

The search is performed using five descriptors or keywords: "recommender system," "blogs" or "blog," "forums or forum," "social network" or "social networks," or "web" or social bookmarking." If opinions between two authors differ, then another author who reviewed the paper is considered to decide whether to include the paper or not. Research papers that are unfit for the review, such as conference submission papers, master and doctoral dissertations, unpublished research papers, non-English papers, survey research papers, and books, are excluded in the review. Research on social media RSs is relatively current; thus, only articles published from 2011 to 2015 in journals related to computer science and information technology from available list of categories are selected to get the highly relevant articles. A total of 47 articles are finally filtered from WOS and 14 articles from Scopus. The selected journals used social media resources, such as digital library, e-commerce, entertainment, forum, geolocation, social network, social bookmarking, and social review.

The five-year period is considered to be representative of the research on social media RSs. Only research papers that describe how RSs can be applied are chosen. A total of 61 research papers are prudently reviewed and classified in terms of the application field and data mining technique. Although the investigation is not exhaustive, it provides a comprehensive basis for understanding social media RSs. The overall graphical classification framework for the selection of research papers is presented in Figure 2.

III. BIBILIOMETRIC ANALYSIS

The bibliometric analysis of the selected 61 studies for the review is conducted. Bibliometric analysis refers to a series of procedures for evaluating a scientific production based on the number of publications, prestige of journals in which the papers are published, and some citations for the publications [77]. Bibliometric refers to a quantitative evaluation of publication and citation data and is currently used in nearly all countries and fields of study. The first idea on quantitative analysis of scholarly output was introduced by Kim and Park [5]. Recent bibliometric studies include research productivity [12], top cited publications [13], and higher education evaluation in a country [14], [15], international collaborations [16], [17], authorship rate [18], and impact assessment of funding environments [19]. Bibliometric analysis of the literature on social media RSs can reveal information on authorship, document type, keyword, and publication trend. Such analysis can also reveal the total number of publications, publications per year, citations per year, highly cited articles, most active author with most published articles, most active country and, most suitable journal.

This analysis is important in the social media RS field and provides results for understanding the number of online social media resources with its rapid growth and difficulty in finding the items in particular specialization by experts.

Therefore, this section uncovers some bibliometric analysis of selected 61 articles. This analysis presents the publication count, citation count each year, top countries contributing in the area of social media in RS, top journals publishing the work on social media in RS, and citation count of each selected article for this review.

Table 1 shows the yearly distribution of the selected articles in both the databases (i.e. WOS and Scopus). Moreover, it also shows the year-wise citation count of publications each year. Here, a fluctuating trend is observed in year-wise publication count with the highest number of publications was published in year 2014, followed by 2015, and 2011 including both the databases. Moreover, the lowest publication count was recorded in year 2012, followed by 2013. The decrease in the number of research papers between 2012 and 2013 is due to the emergence of new RS application fields. The majority of RS studies between 2013 and 2015 are limited to different domains, such as entertainment followed by digital library and social network. A decrease in the publication trend for this research domain is depicted between 2010 and 2011; thus, no publication is found in the later year. Finally, the increasing trend is observed in citation count from year 2011 to 2015. It can be noticed that the citation counts of publications on social media in RS increased from nine (9) in year 2011 to two hundred and seventeen (217) in year 2015. This shows this field is gaining attention of researchers these days.

TABLE 1. Distribution based on year wise citation.

Year	Web of Science		Scopus	
	Articles	Citations	Articles	Citations
2011	10	1	2	8
2012	6	40	2	34
2013	8	71	3	76
2014	11	111	6	112
2015	12	135	1	82

TABLE 2. Distribution based on country wise.

Country Name	Publication Count
South Korea	9
China	9
Iran	8
Spain	8
USA	7
Australia	6

Table 2 shows the top countries contributing in the area of social media RS research. Here, it can be seen that South Korea, and China significantly outnumber all other countries with nine publications each, followed by Iran, and Spain with eight (8) publications each. USA and Australia published seven (7), and six (6) articles each in these five years respectively. Moreover, Canada, and United Kingdom published only three (3) articles each in all five years. The lowest number of articles were published by Malaysia, Greece, Netherlands, Mexico, India, and Argentina.

Table 3 shows the list of top journals published the work on Social Media RS. As shown here, 'Expert Systems with Applications' and 'Information Sciences' are the top journals that published highest number of articles eight (8), and (three (3) respectively in these five years. The rest of the journal published either 2 or 1 articles in all these five years.

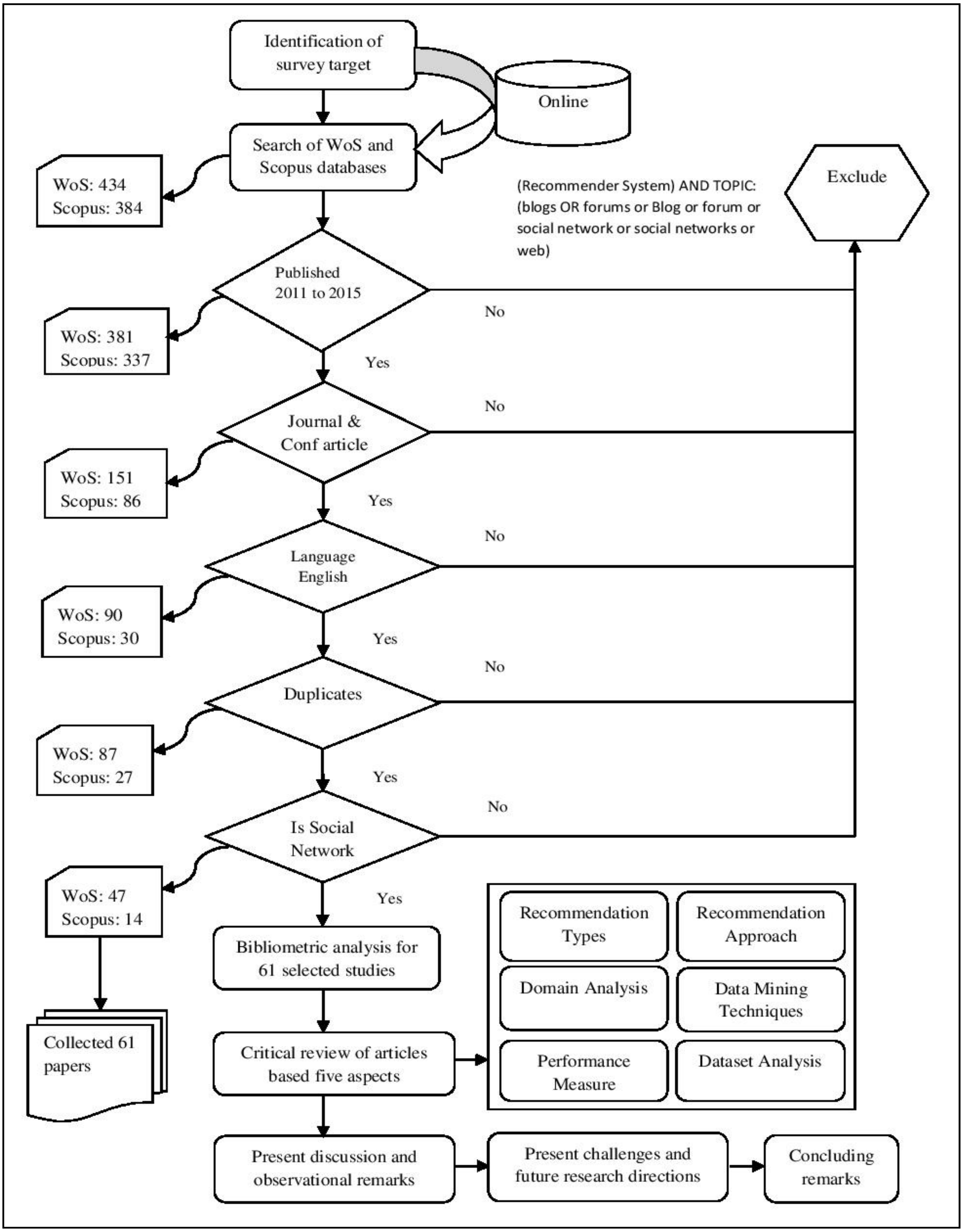


FIGURE 2. Research methodology.

This concludes that for submitting the research work on social media in RS, 'Expert Systems with Applications' is the best choice to submit, followed by 'Information Sciences'.

Table 4 shows the article references obtained from WOS and Scopus with the number of article citations. As can be see here, most of the articles have good number of citation

TABLE 3. Distribution based on journal wise.

Database	Journal Name	Publication Count
Web of Science	Expert Systems with Applications	8
	Information Sciences	3
	User Modelling and User Adapted Interaction	2
	IEEE Transactions on Learning Technologies	2
	Applied Intelligence	2
	ACM Transactions on Information Systems	2
	Others	1 (each)
Scopus	International Journal of Multimedia and Ubiquitous Engineering	2
	Others	1 (each)

TABLE 4. Total number of citations for the 61 selected studies.

S. No	Ref	Citations	S. No	Ref	Citations
1	[4]	5	32	[67]	23
2	[5]	14	33	[8]	2
3	[9]	4	34	[10]	17
4	[11]	1	35	[12]	0
5	[90]	1	36	[13]	1
6	[88]	9	37	[15]	14
7	[16]	55	38	[60]	9
8	[17]	59	39	[18]	0
9	[19]	1	40	[20]	1
10	[21]	0	41	[71]	15
11	[23]	6	42	[22]	3
12	[24]	1	43	[25]	0
13	[32]	6	44	[26]	0
14	[27]	8	45	[87]	41
15	[28]	1	46	[29]	4
16	[30]	0	47	[31]	6
17	[33]	56	48	[61]	1
18	[34]	26	49	[35]	1
19	[36]	33	50	[37]	1
20	[38]	0	51	[39]	1
21	[40]	4	52	[41]	2
22	[42]	4	53	[43]	11
23	[44]	3	54	[45]	6
24	[46]	1	55	[47]	0
25	[57]	8	56	[48]	3
26	[49]	7	57	[59]	26
27	[50]	0	58	[51]	13
28	[52]	0	59	[58]	33
29	[53]	2	60	[54]	0
30	[55]	5	61	[68]	4
31	[56]	2			

count. This shows the quality of articles, selected for this review.

IV. REVIEW AND ANALYSIS

This review identifies six classifications of RS in social media based on its wider usability in recommendation approaches which are, "domains, data mining techniques, datasets, recommendation types, and performance metrics". This analysis can help researchers to gain useful knowledge and better understanding on the existing solutions of social media in RS. The review of all rationale aspects is presented from Section 4.1 to 4.6. This section presents a detailed analysis of the classification for the selected articles.

A. CLASSIFICATION BASED ON RECOMMENDATION APPROACH

In selected articles, several social media RSs employed four common recommendation approaches namely, Content based filtering (CB), Collaborative filtering (CF), Hybrid filtering (HB), and Knowledge based (KB) filtering (as shown in Figure 3).

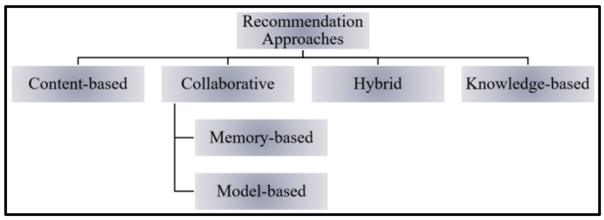


FIGURE 3. Classification based on recommendation.

In selected primary studies, CB technique was extensively used to filter the information in 3 out of 61 articles.

CB methods are used to build several types of representation of the content either by comments or feedback, which are given by the user [13]. Filtering approach recommends the items based on the past choices of the users. This approach generates recommendations using the content of the social media objects used for recommendation such as music, image, sound, or any other products. Similarity can be established between objects based on what the user visited, bought, heard, and ranked. CB filtering has difficulty in identifying user preferences from user actions regarding one content source by using them across all other content types. The recommended content of the same type is limited because the user is already using the value from the RS in social media.

CB filtering is to recommend similar items liked by a user in the past. Items are represented by attributes. Moreover, the recommendation process included both item descriptions and corresponding ratings provided by a user for expressing his/her level of interest. Textual descriptions of item attributes structured into several attributes in several domains. Example domains are movies, music and books which are available as text documents. Attributes such as artist, title, and descriptions are main characteristics for short abstract summarizing of that item using Folksonomy-based Item Recommender System (FIRST). FIRST is found on machine learning techniques for inferring user profiles from both item descriptions and feedback provided as numerical rating. A few studies applying this technique to the user's content for recommendation. The reason for the performance of recommender system improved with folksonomies that might be the valuable source of information in digital artwork recommendation which uses the machine learning techniques. However the authors did not explore the accuracy of impact which improved the accuracy using tagging activity [59].

Different methods such as TAG.ME, Explicit Semantic Analysis (ESA) were used for representing user interests on short text for recommendation. Outcome is the keyword representation used especially on documents like encyclopedia resources. Furthermore, the advantage is linking items and profile, but different sources cannot be merged [11]. In network construction, same interest users are collected as community. Each user named as peer of the network. Links between peers are established between those who have interest on the same content in the past. Network of peers based on shared interest used to identify the recommending item. This limits the public communities from newly joining into peer community with similarity for community building. Hence, companion of the peer structure is not improved [42].

In selected primary studies, 18 out of 61 studies employed CF approach for recommending items. It recommends the items similar to the past choices of users with similar taste that preferences and implicitly requires user ratings.

It is classified into memory-based CF and model-based CF. Memory-based approach is used to compute the rating based on similarity between the users-items. Such approaches are easy to implement because in such approaches the matrices are aggregated to identify the resultant set of items to be recommended [33], [46]. It is flexible with new data and content independent of the items which are being recommended. In model-based approach, data mining techniques are used to develop the prediction models to predict the users rating of unrated items using machine learning algorithms.

Hybrid approach is the most widely used in social media RS. In selected articles, 41 out of 61 have been employed HB approach. This method is a combination of CB and CF as it combines the advantages of both approaches. In recent years, Social media in RS technologies has allowed the emergence of CB and CF to obtain strength points of both the techniques. For example, in social network HB approach is used to combine the rating/feature data of user/item profile to provide an option to overcome the issues of CB and CF [20]. For instance, the authors have employed HB to focus on individual customer behavior or interests [33]. HB exploits individual measures of perceived relevance computed by each user's interest for obtaining better recommendation [37].

In selected articles, 2 out of 61 have been employed KB filtering approach. KB filtering is a specific type of RS that provides explicit knowledge about the items, user preferences, and recommendation criteria [39]. It is advantageous due to the non-existence of cold-start problems. It suggests exactly similar items that were liked by the users in the past. The recommendation for the items semantically related to users' preference. Hence, user preference is not included for recommendation [14].

Figure 4 shows recommendation approach wise and year wise distribution of articles that were used in social media. Furthermore, the pros, and cons of each recommendation approach are shown in Table 5. Finally, the recommendation approach wise and years wise distribution of articles that were used in social media are also shown in Table 6.

B. CLASSIFICATION BASED ON DOMAIN

The selected articles are grouped according to the domains, such as digital library, e-commerce, entertainment, forum, geolocation, social bookmark and social network. The "Others" domain refers to different generic domains (such as, share, publish, and discuss).

Figure 5 shows the classification based on domain of 61 selected articles, related to social media in RS. In selected 61 articles, the identified domains were entertainment, social network, social bookmark, digital library, geolocation, e-commerce, online forum, and "Others". The "Others" category includes some of the digital library domains (such as, UTM website, University of Saskatchewan, and California University). These datasets are university websites and private sector data maintained by their own organizations. Among the selected articles, the most frequently used domain was entertainment (28 out of 61 articles), followed by 7 articles used digital library and social network domains. Around 5 to 6 articles used social bookmark and others (share, publish and discuss domains).Out of 61 articles, geolocation domain adopted in 4 articles. The least research is performed on e-commerce, social review and forum domains which is found on 2 articles in this review. Table 6 shows recommendation approaches with domains and datasets with references of the social media in RS.

As observed from Figure 5, most of the researchers have contributed in entertainments domain. This is because, several freely and publicly available datasets are made available by research community in this domain. These include Flixster, IMDB, MovieLens, and others as some private sectors used in selected articles. However, Forum domain used Stack overflow dataset [82]. Some researchers used multiple datasets in social bookmark, social network and social review in selected articles. However, the focus of recommender system research is more on entertainment domain. The distribution of selected articles by domain and publication years is shown in Figure 9. Here, it can be observed that the domains are evenly distributed in publications during 2011. Until 2013 most recommender systems research domains was focused on entertainment, social bookmark, social network and digital library.

However, the focus of recommender systems research has extended not only to entertainment, but also for geolocation, digital library and social bookmark. Beginning from 2013 research in other domains such as social network, social review, e-commerce, travel, tour etc., are increasing in order year by year.

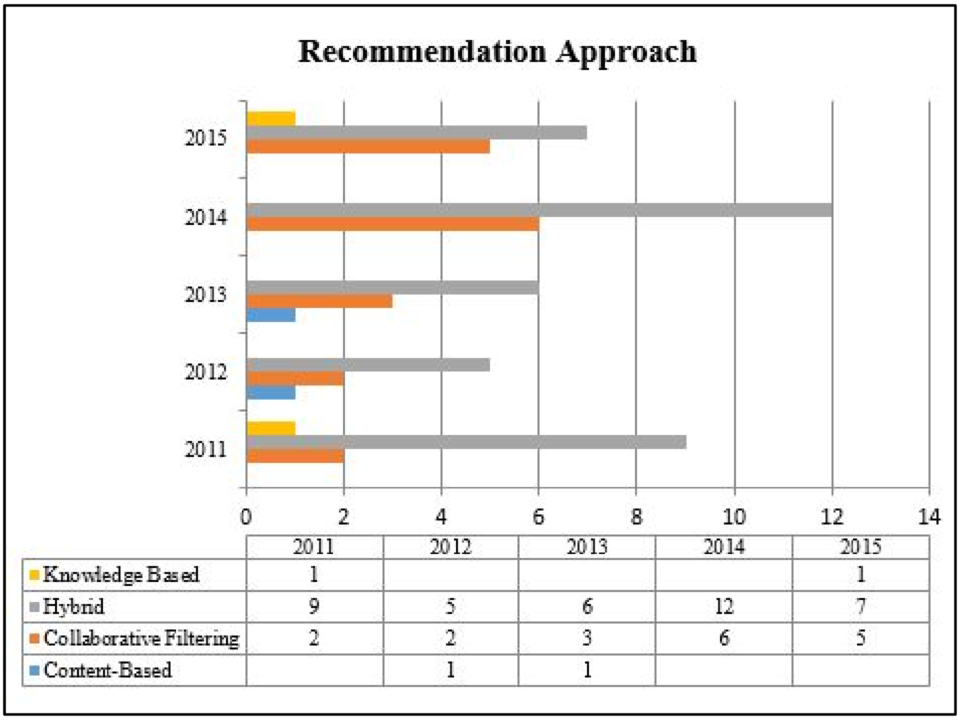
TABLE 5. Pros and cons of RS approaches.

Approaches	Advantages	Main Short comings
CB	CB algorithms explore the text information, which are given by the user as comments/feedback in any domain [10].	Similarity of text is the main issue in audio and video related domain from the larger content.
	Different methods for representing user interests in the form of very short text such as items/users/news titles/feedback. Information retrieval techniques work well in extracting features from text documents. It is adequate to gather comments/feedback from users in the precedence [68].	Practical difficulties in assigning attributes due to limitation of resource.
	Several methods are available to enable an automated creation of communities made from users sharing common interests in past, which is used to build the communities in a simple way.	Limit to public communities. Over-specialization limits the user is to be recommended if the score is high against the user. A new user, having very few ratings, would not be able to give resultant accurate recommendations.
	Personal tags representations are source of information about user interests when combined with the item descriptions.	Limited to users' availability. Items are indistinguishable if items represented by the same set of features.
CF	Providing a list of recommendation based on users' previous interactions between users and items. It is exploring a set of users who have a history of agreeing with the target user by rating the similar set of items.	It is limit to user's preferences from the ratings for new user.
	Calculate the similarity between users' interest of each friend, which can improve the accuracy of recommendation. Clustering the suitable group of friends to calculate the pair-wise user similarities and the similarities among users and items [87].	Limit to new item recommendation which is rely solely on users' preferences. It leads cold start issues [8].
HB	The hybrid system explored more approaches to promote mutual synergies and improve effectiveness and efficiency of the recommendation process.	Increased complexity.
	Combination of the quality of an item with the user preferences which are used to predict the relevancy for generating more useful and accurate recommendations [48].	Small amount of information lead inaccuracies.
KB	Enabling to discover extra knowledge about the user's preferences. Leading more accurate and diverse suggestions which are given by the user [39].	Limited on particular user profile. It is not taken accountable to the behavior of other users.

C. CLASSIFICATION BASED ON DATASET

Various publicly available datasets are available for researchers on social media in RS. The datasets used in the selected 61 articles are examined and summarised in this section. Figure 6 shows the datasets which are used in social media in RS in selected articles.

Table 6 shows the references for datasets and domains. Notably, most of the research papers use more than one dataset termed as "Multiple Dataset" (21 out of 61). For example, datasets from Delicious.com and Movie Lens are used in Tag aware RSs [8]. The "Others" and "Custom" category includes the private network for university data, business transaction data website, some entertainment sectors, travel, television programs and small private sectors, which are maintained by their own organizations. Movie Lens is the mostly used publicly available dataset (12 out of 61 articles). Delicious and Epinions are used in 2 articles out of 61. However, Facebook, Flickr, Stack Overflow, Google wave, Wikipedia and Flixster are the least used datasets of RS in social media. Therefore, researchers on social media in RS almost contributed in all domains, however, very less research has been carried out in academic and CQA related



	2011	2012	2013	2014	2015
Knowledge Based	1				1
Hybrid	9	5	6	12	7
Collaborative Filtering	2	2	3	6	5
Content-Based		1	1		

FIGURE 4. Classification based on recommendation approach.

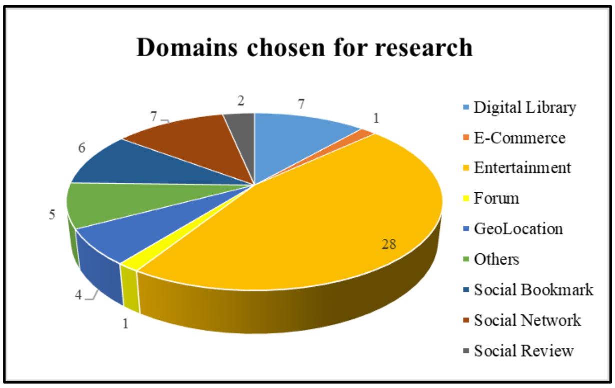


FIGURE 5. Classification based on domain.

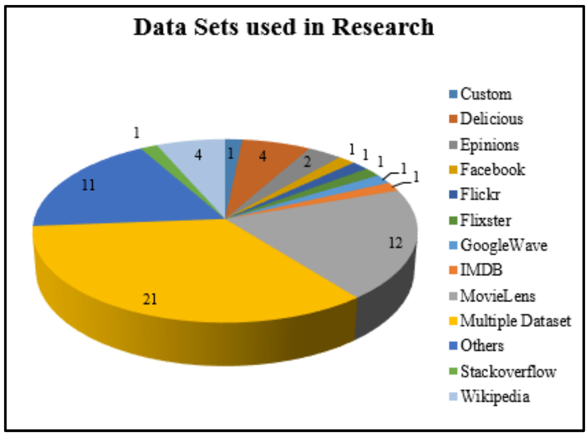


FIGURE 6. Classification based on dataset.

domains during 2011. Important datasets in social media such as Stack overflow and Wikipedia are used by researchers and students. Later on many number of researches used datasets of digital library, social bookmark and forum since 2014.

D. CLASSIFICATION BASED ON DATA MINING TECHNIQUES

Various techniques are used in social media RS to explore and analyze large quantities of data in the form of patterns and rules [4]. These techniques can be used to make a decision on social media in RSs and the effects of decisions are predicted. Data mining techniques were used by many researchers to improve the performance of RS in social media. Figure 7 shows the data mining techniques of social media in RS.

In selected 61 articles, Bayesian Network, and logistic regression techniques were used for CB filtering. Among the data mining techniques the most widely used for CF are clustering, kNN, matrix factorization, link analysis, decision tree, and association rule. Predictions/ranking are then computed within each group. Partitioning or grouping can improve the quality of CF predictions and increase accuracy of performance in RS.

Clustering, matrix factorization, kNN techniques, fuzzy techniques are mostly used in HB filtering. In most of the articles clustering method is used in research articles for grouping the data. Most popular are K-means used especially for big data. From 2011 to 2013 less number of articles used

TABLE 6. References for recommendation approaches, domains and datasets.

Recommendation Approaches	Domains	Datasets	References
CB	Digital Library	Others	[59]
	Social Network	Multiple Datasets	[13]
	Digital Library	Others	[46] [53]
	E-Commerce	Others	[34]
CF	Entertainment	MovieLens	[31][33][25][50][83]
		Epinions	[56]
		Multiple Dataset	[48][18]
	Others	Wikipedia	[10]
	Social Bookmark	Delicious	[52][37][68]
	Social Network	Multiple Dataset	[30][47]
	Social Review	Epinions	[60]
	Digital Library	GoogleWave	[35]
		Others	[44][41][43]
HB	Entertainment	Flixster	[26]
		IMDB	[12]
		MovieLens	[36][16][5][20][84][27]
		Multiple Dataset	[23][49][32][54][61]
		Others	[40][24][21]
	Forum	Stack Overflow	[38]
	GeoLocation	Custom	[11][6]
		Multiple Dataset	[58][15][28]
	Others	Flickr	[19]
		Multiple Dataset	[86]
		Wikipedia	[9][57][25]
	Social Bookmark	Delicious	[68]
		Multiple Dataset	[67][8]
	Social Network	Facebook	[55]
		Multiple Dataset	[51][4][45]
	Social Review	Multiple Dataset	[22]
KB	Entertainment	Multiple Dataset	[39]
		Others	[17]

clustering technique. Since 2013, most of the researchers proposed clustering-based techniques (shown in Figure 5). Compared to kNN, the performance of clustering is improved when the instances are close to the boundaries for large size data. However, kNN explored the poor efficiency of using instance-based method. Other techniques such as link analysis, logistic regression and trust techniques are least used in the selected articles in social media RS from year 2011 to 2015. The commonly used techniques in the selected studies are discussed below.

1) CLUSTERING

Clustering algorithms are used mainly to partition the set of items based on user rating data. The clustering method identifies a finite set of categories or clusters to describe data. This technique is applicable for large dataset [34]. Among the clustering methods, "K-means clustering" is the most popular. K-means considers the input parameter K and partitions a set of n objects into K clusters. Clustering is scalable, simple, and suitable for datasets with compact and highly distributed spherical clusters. Clustering or classification techniques are increasingly applied to recommendation methods in social media to enhance the recommendation accuracy [44]. Out of 61 articles, 17 articles used clustering technique most frequently. The users in a cluster can have similar interests; if a product is selected by these users, then this product can be suitable for the target user. This way leads to more accuracy in recommendations for social media in RS.

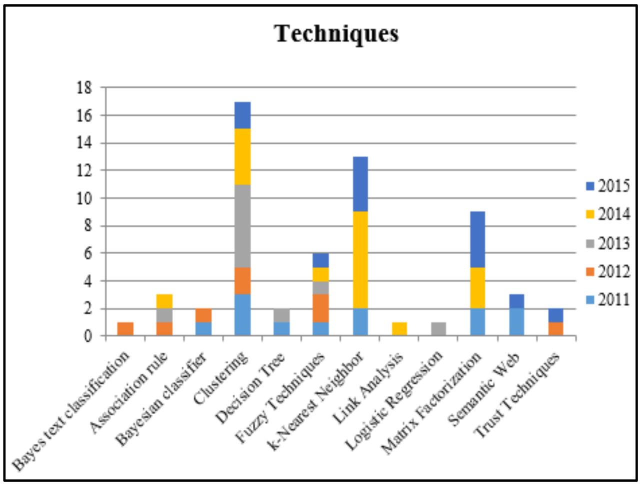


FIGURE 7. Classification based on data mining technique.

Highlighted advantage is the good performance for using clustering techniques in very large scale problems in recommender system which is used to work on aggregated data [6]. Several cluster based approaches reduces computation time. Approaches are used to group the users' preferences in order to increase neighbor searching efficiency. Disadvantage is limitation of clustering techniques which often lead to poor prediction accuracy than other methods. Finally, high dimensional dataset lead issues in clustering [37].

2) K-NEAREST NEIGHBOR (*k*NN)

Followed by clustering, *k*NN is used in 13 articles out of 61. The *k*NN model is a typical traditional CF-based social media in RS that makes recommendations according to the following three steps. (1) RSs construct a user profile using user preference ratings. (2) RSs apply statistical or machine learning techniques to discover k users (i.e., neighbors or recommenders) who exhibit similar behaviors in the past from which a neighborhood is formed based on the degree of similarity between a marked user and the other users. (3) Once a neighborhood is formed for a target user, RSs make a top n items set for that target user who is most likely to purchase by analyzing the items interested by the neighbors [50]. The effectiveness of a representation based on user profiles and items are then used in the investigation. Items and concepts are associated to user profiles, and this kind of information is useful for generating readable user profiles [20].

The strengths of *k*NN are: it is very simple algorithms to implement, moreover it is easy to explain, understand and interpret. In addition, the accuracy predicted by *k*NN is good enough. Flexible features are generally used to handle multi-class cases and practicing with enough representative data which were mostly able to adopt for classification or regression. A *k*NN classifier is memory-based, which is not requiring any model to express in simple way [26]. The set of observations apart from collecting vectors of features with labels are enough for classes which they belong to items. The iterative clustering approach allows them to solve the data sparseness problem by fully exploiting the voting information first followed by applying cluster-based *k*NN to improve the performance [47].

The accuracy can be compromised and affected by noise or irrelevant features. The memory constraints can be applied for large search problem to find k nearest neighbors. It is computationally very expensive due to high memory requirement to store all training data in the memory. Prediction stage might be slow. It is sensitive to irrelevant features and the scale of the data [46].

3) MATRIX FACTORIZATION

Matrix factorization is the data mining technique next to *k*NN, which is used in 9 articles out of 61. The characterized items and users grouped by vectors of factors inferred from item rating patterns. Various input data are collected and placed in a matrix representing users and other dimension representing items of interest. The most convenient data are high-quality explicit feedback, which includes explicit input by users about their interest in products [48]. Matrix factorization allows incorporation of additional information. When explicit information is unavailable, the RS in social media can infer user preferences using implicit feedback, which indirectly reflects user opinions. Implicit feedback usually denotes the presence or absence of an event and is thus typically represented by a densely filled matrix. Main pros of matrix factorization characterize both items and users by vectors which are based on item rating. High correspondence between item and user factors leads accuracy in recommendation, by combining scalability [15]. Moreover, strength is very flexible for modeling in various real time situations. Strength of matrix factorization is that it allows incorporation of additional information. The major issue is memory management with large-scale and computational effort. High memory utilization for more data to be processed and more CPU time is needed.

4) FUZZY TECHNIQUES

Next to matrix factorization technique fuzzy technique is used in 6 articles out of 61 articles. Fuzzy technique defines the items in many systems by using specified rating scale. For example, rating on a scale of 1 to 5, where 1 indicates the lowest preference and 5 indicates the highest preference for an item by a specific user. Some systems also have various preferred models in specific application fields. In such systems, users express their preferences in linguistic terms, such as "interested" or "not interested" or "like" or "dislike." The recommendations for these systems are generated on the basis of uncertain of vague information. The similarities between items or between users are fuzzy; thus, many researchers apply fuzzy set theory, fuzzy logic, and fuzzy relations to RS in social media to achieve accurate and effective recommendations [35].

5) SEMANTIC ANALYSIS

Semantic analysis technique is next to fuzzy technique which is used in 3 articles out of 61. This is for extracting and

TABLE 7. Data mining techniques used in selected studies.

Techniques	References
Bayes text classification	[59]
Association rule	[52][11][87]
Bayesian classifier	[32][12]
Clustering	[40][42][44][9][58][8][10][5] [55][24][31][21][34][37][26] [61][22][71]
Decision Tree	[47][53]
Fuzzy Techniques	[35][36][38][33][43][45]
k-Nearest Neighbor	[46][16][20][27][28][19][23] [41][49][50][29][66][86]
Link Analysis	[30]
Logistic Regression	[13]
Matrix Factorization	[48][51][54][15][4][18] [25][67][84]
Semantic Web	[57][17][39]
Trust Techniques	[60][56]

representing the contextual-usage meaning of words by statistical computations applied to a large corpus of text [17]. The underlying idea is to aggregate all word contexts, for example, set of words "does or does not" which provide a set of mutual constraints that largely determines the similarity of meaning of words by sematic process. This explores to analyze for creating user profiles from the contents by examining posts/shared information.

Other techniques such as association rule, Bayesian classifier, link analysis, decision tree and trust techniques are also used in social media RS. Hence, clustering technique is highly active in several fields such as, image classification, statistics, text classification, pattern recognition, and machine learning.

The clustering approaches are more popular because these techniques can be exploited on large datasets with many attributes in it. This imposes unique computational requirements on relevant clustering algorithms for grouping the data. Variety of algorithms has recently emerged that meets these requirements. Moreover, clustering techniques were successfully applied to real-time data mining problems in variety of datasets. Table 7 shows the study-wise, and recommendation approach-wise distribution of data mining techniques in social media in RS field.

E. CLASSIFICATION BASED ON RECOMMENDATION TYPE

RS in social media provide recommendation based on two approaches. In selected 61 articles, two approaches are used

1. Classical or two dimensional (2D).
2. Contextual or multidimensional approach.

Generally, classical approaches fall under three recommendation approaches which are CB, CF and HB. It includes recommendation types such as, Items X User, User X Item, User X Tag, and User X User. In most of the research articles, ratings (R) are based on two of entities, user and item.

R = Item x User Relevance between pair of item and user
R = User $\times$ Item, Relevance between pair of user and item.
R = User $\times$ *Tag*, Relevance between a pair of user and tag.
R = User $\times$ User, Relevance between a pair of user and user.

For the two-dimensional approach,
R = User $\times$ Item or User $\times$ Tag or User $\times$ User.

Most of the articles [27], [56], [18], [23] use classical approach/2D in terms of user-item, item-user, user-tag and user-user. CB filtering recommendation approach explored between content of items and CF explored between users. HB is to reduce the limitations and improve the efficiency. KB employed between user and items.

Contextual/Multi-dimensional approaches utilized dimensions which includes User $\times$ Tag $\times$ Items or User $\times$ Item $\times$ Tag.

Thus, the resultant rating of multi-dimensional approach represents

$$R = D_1 \times D_2 \times D_3 \times D_4 \times \ldots\ldots\ldots \times D_n.$$

Considerable research has been carried out in social media RS. Commonly used recommendation type is user - item. Moreover, the least used recommendation type is user-tag recommendation type followed by user-user. Table 8 shows the recommendation types used in each study. Few studies [35] have used contextual approach.

This approach used more than one attribute for recommendation, for example entities in digital library domain (book id, author, subject and year). Hence, recommendation function will become multiple attributes or multi-dimensional in the matrix format leads more accuracy in recommendation.

F. CLASSIFICATION BASED ON PERFORMANCE METRICS

In selected articles, researchers utilized precision, recall, accuracy, weight, mean absolute error, and efficiency. In most of the studies, authors have used either precision, recall, and accuracy metrics or only accuracy to measure the performance of RS (as shown in Table 9). Precision and recall evaluation metric in the selected articles [3], [22] was based on the quality of similarities obtained among users/items. Obtained resultant similarities are more reliable, to lead the better performance for RS in social media. Various studies in the selected articles utilizes all user activities stored in separate layers and evaluated by weight metric [16]. Authors employed Mean Absolute Error (MAE), a metric measure based on correlation between predictions and ratings [83]. Accuracy metric mainly measure the comparison of the estimated ratings against actual ratings in the User Item. In research [15], researchers utilized accuracy measure by statistical data of ratings.

V. DISCUSSION

In this review, 61 articles regarding social media RSs were selected from WOS and Scopus. Among these papers, 47 were selected from WOS, and 14 were obtained from

TABLE 8. Recommendation type employed in selected primary studies.

Study	IT	UI	UT	UU	Study	IT	UI	UT	UU	Study	IT	UI	UT	UU
[4]	✗	✓	✗	✗	[29]	✗	✗	✓	✗	[49]	✗	✓	✗	✗
[5]	✗	✓	✗	✗	[30]	✗	✓	✗	✗	[50]	✗	✗	✓	✗
[8]	✗	✓	✗	✗	[31]	✗	✓	✗	✗	[51]	✗	✓	✗	✗
[9]	✗	✓	✗	✗	[32]	✗	✓	✗	✗	[52]	✗	✗	✓	✗
[10]	✗	✓	✗	✗	[33]	✗	✓	✗	✗	[53]	✗	✓	✗	✗
[11]	✗	✓	✗	✗	[34]	✗	✓	✗	✗	[54]	✗	✓	✗	✗
[12]	✗	✓	✗	✗	[35]	✗	✓	✗	✗	[55]	✗	✓	✗	✗
[13]	✗	✓	✗	✗	[36]	✗	✓	✗	✗	[56]	✗	✓	✗	✗
[15]	✗	✓	✗	✗	[37]	✗	✓	✗	✗	[57]	✗	✓	✗	✗
[16]	✗	✓	✗	✗	[38]	✗	✗	✓	✗	[58]	✗	✓	✗	✗
[17]	✗	✓	✗	✗	[39]	✗	✓	✗	✗	[59]	✓	✗	✗	✗
[18]	✗	✓	✗	✗	[40]	✗	✓	✗	✗	[60]	✗	✗	✗	✓
[19]	✗	✗	✓	✗	[41]	✗	✓	✗	✗	[61]	✗	✓	✗	✗
[20]	✗	✓	✗	✗	[42]	✗	✓	✗	✗	[67]	✗	✓	✗	✗
[21]	✗	✓	✗	✗	[43]	✗	✓	✗	✗	[68]	✗	✗	✓	✗
[23]	✗	✓	✗	✗	[44]	✗	✓	✗	✗	[22]	✗	✓	✗	✗
[24]	✗	✓	✗	✗	[45]	✗	✓	✗	✗	[84]	✗	✓	✗	✗
[25]	✗	✓	✗	✗	[46]	✗	✓	✗	✗	[86]	✗	✓	✗	✗
[26]	✗	✗	✗	✓	[47]	✗	✓	✗	✗	[87]	✗	✓	✗	✗
[27]	✗	✓	✗	✗	[48]	✗	✓	✗	✗	[71]	✗	✓	✗	✗
[28]	✗	✓	✗	✗										

****IT** = Item–User, **UI** = User–Item, **UT** = User–Tag, **UU** = User–User

Scopus. All the 61 articles were selected using various filters, such as language, year, and category.

In the bibliometric analysis, articles from WOS and Scopus were classified according to years, countries, and journals. Articles about social media RS are only few in 2010, but the number increases annually. After bibliometric analysis, the filtered articles were classified on the basis of recommendation type, domain, data mining technique, dataset, and performance metric. This study on social media RSs revealed various notable information.

A. RECOMMENDATION APPROACHES

HB, CF, and CB filtering are the most used recommendation types in sequence between 2011 and 2015 [49], [56]. Various types of recommendation filtering display distinct advantages and disadvantages. CB filtering makes recommendations similar to items previously preferred by a specific user. These items are used to distinguish the items used in user profile. To compare each item with the user profile, items with a high degree of similarity with the user profile will be recommended [6]. CF is used to select based on the opinion of other people who share similar interests. CF can be classified into user-based and item-based CF approaches.

In user-based CF approach, a user will receive an item recommendation liked by similar currently active users. In an item-based CF approach, the user will receive recommendations of items by those who liked the item in the past [3], [9]. A prediction for the active user is calculated by a weighted average of the ratings of the selected users. The important issue with CF is sparsity; having only few ratings is a considerably severe problem [31]. HB achieves high performance and overcomes the drawbacks of CB and CF [55]. To overcome the existing issues on CB and CF, HB approach is used in common practice to avoid cold start, sparseness, and/or scalability problems [34] and obtain improved performance [75]. HB is mostly custom-based development. Different social media recommendation types and strategies are combined and exploited under HB to improve the performance of recommendations when individual techniques do not provide satisfying results. Several studies showed that HB approach achieves good performance in RS [10], [50], [58] by combining the advantages of CB and CF approaches.

B. RS DOMAINS

Review analysis indicated that social media RSs have been applied in various domains, such as digital libraries, e-commerce, entertainment, geolocations, forums, social networks, social reviews, social bookmarks, and other (share, publish, and discuss) domains. With the extensive usage of entertainment domains in recent years, a particularly rapid growth in movie, video chat communications, multiplayer gaming, music and video streaming, live video streaming and music resources has occurred [47]. Entertainment datasets

TABLE 9. Performance metrics used in selected primary studies.

Study	AC	EF	MAE	PR	RC	WT	Study	AC	EF	MAE	PR	RC	WT
[4]	✗	✗	✗	✓	✗	✗	[39]	✗	✗	✗	✗	✓	✗
[5]	✗	✗	✗	✗	✓	✓	[40]	✗	✗	✗	✗	✓	✗
[8]	✓	✗	✗	✗	✗	✗	[41]	✗	✗	✗	✗	✓	✗
[9]	✓	✗	✗	✗	✗	✗	[42]	✗	✗	✗	✗	✓	✗
[10]	✓	✗	✗	✗	✗	✗	[43]	✗	✗	✗	✓	✗	✗
[11]	✗	✓	✗	✗	✗	✗	[44]	✗	✗	✗	✗	✓	✗
[12]	✓	✗	✗	✗	✗	✗	[45]	✗	✗	✗	✗	✗	✓
[13]	✓	✗	✗	✗	✓	✗	[46]	✗	✗	✗	✗	✗	✓
[15]	✗	✗	✗	✗	✓	✗	[47]	✗	✗	✗	✗	✓	✗
[16]	✗	✗	✗	✗	✗	✗	[48]	✓	✗	✗	✗	✗	✗
[17]	✓	✗	✗	✗	✗	✗	[49]	✗	✗	✗	✗	✗	✓
[18]	✗	✗	✓	✗	✗	✗	[50]	✗	✗	✗	✗	✗	✓
[19]	✗	✗	✗	✗	✗	✓	[51]	✓	✗	✗	✗	✗	✗
[20]	✗	✗	✗	✗	✗	✓	[52]	✓	✗	✗	✗	✗	✗
[21]	✗	✗	✓	✗	✗	✗	[53]	✗	✗	✗	✓	✗	✗
[23]	✓	✗	✗	✗	✗	✗	[54]	✓	✗	✗	✗	✗	✗
[24]	✓	✗	✗	✗	✗	✗	[55]	✗	✗	✗	✗	✓	✗
[25]	✗	✗	✗	✗	✓	✗	[56]	✗	✗	✗	✓	✗	✗
[26]	✗	✗	✗	✗	✗	✗	[57]	✗	✗	✗	✗	✓	✗
[27]	✗	✗	✗	✗	✗	✗	[58]	✓	✗	✗	✗	✗	✗
[28]	✗	✗	✗	✗	✗	✗	[59]	✗	✗	✗	✗	✓	✗
[29]	✗	✗	✗	✗	✓	✗	[60]	✗	✗	✗	✓	✗	✗
[30]	✗	✗	✗	✗	✓	✗	[61]	✓	✗	✗	✗	✗	✗
[31]	✗	✗	✗	✗	✓	✗	[67]	✗	✗	✗	✗	✓	✗
[32]	✗	✗	✗	✗	✓	✗	[68]	✗	✗	✗	✗	✓	✗
[33]	✗	✗	✗	✗	✗	✓	[22]	✗	✗	✗	✗	✗	✓
[34]	✗	✗	✗	✗	✓	✗	[84]	✓	✗	✗	✗	✗	✗
[35]	✗	✗	✗	✗	✓	✗	[86]	✗	✗	✗	✗	✗	✓
[36]	✓	✗	✗	✗	✗	✗	[87]	✗	✗	✗	✗	✓	✗
[37]	✗	✗	✗	✗	✓	✗	[71]	✓	✗	✗	✗	✗	✗
[38]	✗	✗	✗	✗	✓	✗							

****AC** = Accuracy, **EF** = Efficiency, **MAE** = Mean Absolute Error, **PR** = Precision, **RC** = Recall, **WT** = Weight

are used in 46% of all the selected papers in RSs. Movie recommendations combined with multimedia platforms are widely used, thereby providing users with preferred movie results that are intelligently filtered and calculated from a large movie database [35]. Most of the research is developed in movie recommendation [6], [44]. Research on digital library domain deals with search engines that calculate the relevance between research papers and search queries, such as research paper recommendation systems. These RSs produce recommendations on research papers using the keywords and interests of users.

Digital libraries and social networks are used individually in 11% of the papers. Social bookmarks are used in 10% of the papers based on the user–item recommendation type [52]. Forum domains use recommendation systems for recommending experts and posts (1%). Forum sites, such as Stack Overflow, display large user participation by posting questions and answers. Expert recommendation for an input question will be helpful to receive quality answers. Posts in forums are archived and used for recommending answers to newly posted similar questions. In e-commerce domain, people are interested in real-time recommendations (2%), where the user's interests or preferences for some items continue to change. To satisfy these requirements, real-context-awareness-based recommendation methods should be used.

Social bookmark domains are mainly used for tagging pages that are stored on the web; these pages can also be accessed from any computer [26]. Geolocation domain, such as travel websites, is used to determine locations and provide details. RSs applied for social bookmark and geolocation are

10% and 7%, respectively. Since 2013, research on social bookmark, social network, forum, and e-commerce domains is increasing, as shown in Figure 9.

C. DATASETS USED IN RS IN SOCIAL MEDIA

The most popular datasets used in social media RS are Movie Lens, IMDB, and Flixster. These datasets are mainly used due to the free data access provided to users. MovieLens is a well-known dataset on movie recommendations used by the research community. This dataset contains the users and the rating given by users to different movies [33]. MovieLens also searches for similar profiles, such as users that share the same or similar taste, and uses them to generate new suggestions. This review evidently indicated that several researchers used this dataset for research and evaluation of their proposed RS technique [15].

MovieLens dataset is used in 12 articles. The least amount of research is performed on academic sites in digital libraries and forums (such as Stack Overflow), and other (share, publish, and discuss) domains. Datasets related to the entertainment domain are mostly used in the RS research [17], [57] for the selected articles.

The MovieLens dataset is frequently downloaded and referenced in research articles. This dataset is highly popular because of its significance in exploring the flexibility in rating data, dataset update (recently MovieLens V4), and updates on validating ideas. This dataset applies recommendation technology to determine recommendation and includes pattern identification, visualization, comments about personal taste, and flexibility in personalization technology, which are important attributes. Recent release in MovieLens included many attributes to encourage the development of educational materials with high quality, software systems, startup companies, and academic research [22]. Other available datasets are used less frequently compared with MovieLens in selected articles. For example, IMDB dataset is the second largest movie collection dataset, but only plain-text data are available for download. The main restriction of this dataset is being publicly unavailable for research purposes. Epinions is another dataset that is used in the selected articles. This dataset contains user–item reviews, which are available in SQL format. Domain independence and rich information make Epinions one of the few public datasets. Delicious dataset is used to store and share social bookmark with tags. This dataset is publicly viewable, but the user can download his/her data only [78]. Wikipedia presents unstructured data with Wikipedia pages. This dataset is available in human-readable form that is less noisy, but data extraction in RS is difficult [11], [85].

Stack Overflow is a forum dataset that can be accessible from the Stack Exchange website. Data are available in XML format with many attributes, such as comments and badges [79]. Other datasets include Facebook, Googlewave and certain custom datasets. Most of entertainment domains are available as open source for download of data with many attributes, such as author, time, rating, comment, free-text tagging, and genre. Genres are pipe-separated lists, such as adventure, documentary, horror, musical, and mystery [80], [81]. Researchers can select the attributes on the basis of their research problem requirements.

D. DATA MINING TECHNIQUES

Review on various data mining techniques for RS showed that kNN, clustering, and matrix factorization are the most widely used techniques in RS. kNN is a traditional CF-based algorithm utilized for various purposes in RSs, such as constructing user profiles with ratings [26]. Social data extractor extracts textual information about user activities, especially on social networks (such as Facebook and Twitter) and user information (such as title and description of liked groups, attended events, liked pages, and articles). Tweet information and direct messages can also be used to generate user profiles [47]. User profiles constructed using these techniques are useful in expert and item recommendations. The recommendation accuracy can be improved by combining various techniques.

Clustering is scalable, simple, and suitable for datasets with compact and highly distributed spherical clusters. Clustering or classification techniques are increasingly applied into recommendation methods to enhance the accuracy of recommendation [31]. Users in a cluster can display similar interests. If a product is selected by these users, then this product can be suitable to the target user. This method results in accurate recommendations. In clustering techniques, a given data set is organized into group of clusters in a manner that data points are similar. This organized cluster is used to provide recommendations on related items. Clustering is helpful in generating recommendation from large datasets. Clustering techniques are divided into two types: hierarchical and partitioned clustering [34]. Hierarchical clustering proceeds successively by either splitting large clusters or merging small clusters into large ones. Partitioned clustering directly breaks down data set into a set of disjoint clusters [57].

Matrix factorization is effective, and it is used to determine the latent feature association between users and items [15]. Matrix factorization creates matrices for the association of users and items and calculates a score. A recommendation is generated on the basis of the score [51]. Other techniques, such as association rule, Bayesian classifier, link analysis, decision tree, and trust techniques, are also used in social media RS in selected articles.

E. TYPE OF RECOMMENDATION

Review on various recommendation types, including item-tags, user-items, user-tags, and user-user, showed that the most used recommendation type in the selected papers for review is user-item [22], [30], whereas the least used type is item-tag [35]. Many recommendation systems recommend items to users based on the generated profile; for example, the MovieLens website recommends movies to users based on their interest. E-commerce websites utilize user–item recommendations based on user's interest, such as electronics and clothing. The interaction between user–item can be implicit or explicit. Implicit interactions include sessions and cookies of browser, whereas explicit interactions are through the ratings and feedback provided by user. Social bookmarking datasets are less used in social media RS research compared with social network datasets. Other social media resources, including forums, blogs, video portals, and image portals, are also used less frequently in existing research articles. Although a large number of datasets are available in social networks, the number of research is small in social networks, including academic, forum, and digital library portals. Tags are arbitrary words specified by authorized users to label and manage contents, which are uploaded to the forum websites by user-tag recommendation type. Item-tag recommendation is used in digital library domains. Social network domains are explored by user-user recommendation type in social media RS. Figure 8 shows the number of publications in domain versus the type of recommendation.

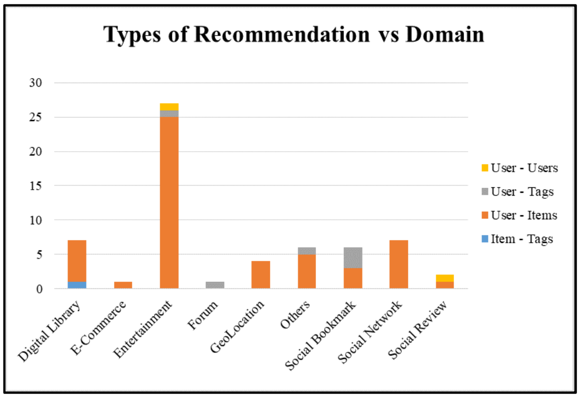


FIGURE 8. Social media research vs. domain.

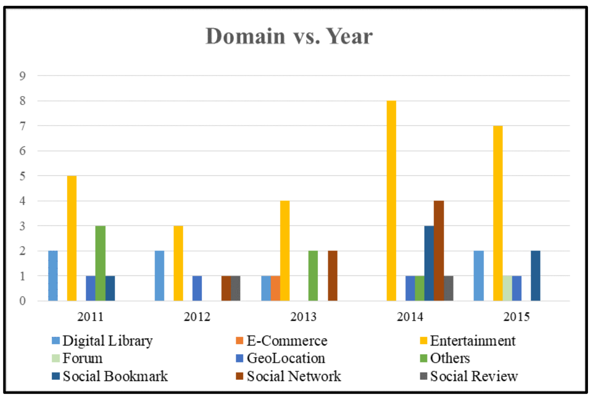


FIGURE 9. Year-wise domain publications.

For effective sharing of content, users want personalized and convenient tags. However, selecting appropriate tags from a wide range of possibilities is often difficult among users. Thus, tag RSs become increasingly important for making tag selection effective mainly for recommendation.

F. PERFORMANCE METRICS

The predictions generated by RS in social media can be evaluated using a variety of performance metrics. RS predictions can be evaluated with prediction accuracy. Other aspects of RS performance are time and space requirements. Response time is the time required to generate responses to the request of users. Main and secondary storage requirements are also evaluated for RSs. These evaluations can be performed online, offline, or both based on the requirement.

In selected studies, various metrics were used to evaluate the performance of RS. The review showed that precision and recall are highly used evaluation metrics; they are important in information retrieval and popular with RS in social media. Precision measures the efficiency of relevant items selected. Recall is the number of relevant items selected and the total number of items in the whole set. Recall is used to access the extent of the result [26]. Although precision and recall are different measures, they cannot be considered unrelated. Accuracy metric is also highly used for evaluation. This metric measures the ranking accuracy of results predicted by a recommendation system. Accuracy metric measures how well the user ratings can be reproduced [49]. Analysis on performance measures can help researchers select the suitable performance metrics for their research. Different evaluation metrics can be used in measuring the performance of social media RS.

VI. POTENTIAL ISSUES

Numerous studies highlighted several significant problems and issues. The discussion demonstrated that different classifications can be used to observe the summarized data in each section.

A. DEFICIENCY IN DOMAINS

A majority of reviewed articles are related to entertainment domains (46%, 28 articles). Studies on forum and social review domains are only 2% (one article) and 3% (two articles), respectively [85]. The number of research in these domains is another significant direction.

B. DEFICIENCY IN DIGITAL LIBRARY RESEARCH

Academic-related digital library websites are mostly owned by private universities as collection of information objects. Academic scholars, research students, and technical people use forums and search engines, such as Google Scholar, Stack Overflow, and Quora. Future research can use academic-related library [3]. Forum datasets are a significant direction for further exploration.

C. E-COMMERCE RECOMMENDATION

In recent years, people increasingly explore online shopping websites. E-commerce websites require a social media RS to determine the proper product recommendation. Distribution of data, such as user behavior and interest and liked items, continues to change in e-commerce websites [26]. Further research may use contextual recommendations, such as history, time, and preference recommendations, for the e-commerce domain [83].

D. USER-TAG TYPE OF RECOMMENDATION

User–tag type of recommendation is used in only a few domains. Social networks and other domains extensively use user–item recommendation type [49]. Tag RS is important in making tag selection for RS in social media for creation of user profile to improve the performance of RS for technical people. Other explorations based on tag type and other attribute relations, especially in forum and digital library datasets, are another direction.

E. DATA SPARSITY

In handling data sparsity problem in RSs, various methods are proposed, but the issue still persists in many application domains. Data from various domains can be combined in target domain to handle sparsity problem. RS approaches are limited with the selected method and data collected from the source [34]. Analysis on RSs can be performed on multiple dimensions using large data for remarkably accurate recommendations. For example, highly accurate recommendations can be made in data sources on social bookmark and social network domain by extracting a large amount of information about people and items.

F. RECOMMENDER SYSTEMS TO ASSIST IN ASSISTING SECURITY IN SOCIAL MEDIA

Social media systems are prevalently growing. Recently, new challenges are integrated with such systems, thereby making this system difficult to manage. New reliable solutions to such problem should be developed. Consequently, important, practical, and analytical solutions to such systems can be proposed to target security assurance in social networks and social human interactions [62], [63]. To address this problem, a recommendation system can also be developed to verify social data management in IoT devices.

VII. CONCLUSION

This comprehensive study on RS in social media is an important research area that has attracted the attention of the academia and practitioners. In this research, 61 papers on social media RSs published between 2011 and 2015 were comprehensively reviewed and examined to understand the trend of these systems.

All major research efforts in this regard assist researchers in gaining an improved understanding regarding the existing solutions to major problem areas. The selected articles were analyzed. Results of this review are summarized as follows:

- Interest in RS of social media research will grow significantly in the future. Yearly distribution shows an increasing trend in the last few years.
- Classification based on domain shows that 19 among 61 papers are related to movies, and 14 papers are related to entertainment. Social media RSs in many other domains require significant improvement.
- The classification based on data mining technique shows that clustering is the mostly used technique. A total of 17 out of 61 papers use this technique for social media recommendation. The least used techniques are link analysis and Bayesian text classification.
- Classification based on dataset shows that most studies use multiple datasets from various datasets. The highly utilized dataset for social media RS is MovieLens because of the provided free data access to users.
- A majority of social media RS studies are published in Experts System with Application. Additional papers on RS research should be published in journals related to management and business.

This study also displays some limitations: (i) this review only includes articles that are limited to computer science and information technology journals. (ii) Reviewed journals are selected from 2011 to 2015 journals that are related to social media. Journals published after 2015 are excluded. This duration can be extended in future reviews. (iii) The limitation on keywords is described as follows: ''recommender system,'' ''blogs'' or ''blog,'' ''forums or forum,'' ''social network'' or ''social networks,'' or ''web'' or social bookmarking.'' (iv) This review is based only on recommendation approaches, data mining techniques, datasets, domains, recommendation types, and performance metrics. Other significant aspects were not analyzed and reviewed.

The present results can contribute significantly in understanding social media in RSs. A total of 61 articles out of 434 were rigorously selected and reviewed. The classification of social media RS research in terms of recommendation approach, domain, data mining technique, dataset, recommendation type, and performance metric can also provide detailed insights into these aspects and guidelines for future RS research in social media.

REFERENCES

[1] A. Tiroshi, T. Kuflik, J. Kay, and B. Kummerfeld, ''Recommender systems and the social Web,'' presented at the Int. Conf. User Modeling, Adaptat., Personalization UMAP, Aug. 2011, pp. 60–70.

[2] A. Tiroshi, T. Kuflik, J. Kay, and B. Kummerfeld, ''Recommender systems and the social Web,'' in *Proc. UMAP 19th Int. Conf. Adv. User Modeling*, Berlin, Germany, 2011, pp. 60–70.

[3] F. Ricci, G. Semeraro, M. de Gemmis, and P. Lops, ''Decision making and recommendation acceptance issues in recommender systems,'' in *Proc. UMAP 19th Int. Conf. Adv. User Modeling*, Berlin, Germany, 2011, pp. 86–91.

[4] M. Diaby, E. Viennet, and T. Launay, ''Exploration of methodologies to improve job recommender systems on social networks,'' *Social Netw. Anal. Mining*, vol. 4, no. 1, p. 227, Dec. 2014.

[5] M. Kim and S. O. Park, ''Group affinity based social trust model for an intelligent movie recommender system,'' *Multimedia Tools Appl.*, vol. 64, no. 2, pp. 505–516, May 2013.

[6] U. Shardanand and P. Maes, ''Social information filtering: Algorithms for automating 'word of mouth,''' presented at the CHI SIGCHI Conf. Hum. Factors Comput. Syst., May 1995.

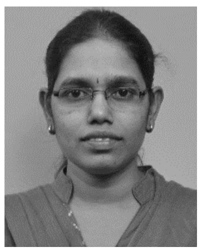
[7] E. Garfield, ''Citation indexes for science: A new dimension in documentation through association of ideas,'' *Science*, vol. 122, no. 3519, pp. 108–111, Jul. 1955.

[8] Y. Song, L. Zhang, and C. L. Giles, ''Automatic tag recommendation algorithms for social recommender systems,'' *ACM Trans. Web*, vol. 5, no. 1, p. 4, Feb. 2011.

[9] E. Davoodi, K. Kianmehr, and M. Afsharchi, ''A semantic social network-based expert recommender system,'' *Appl. Intell.*, vol. 39, no. 1, pp. 1–13, Jul. 2013.

[10] B. Amini, R. Ibrahim, M. S. Othman, and A. Selamat, "Capturing scholar's knowledge from heterogeneous resources for profiling in recommender systems," *Expert Syst. Appl.*, vol. 41, no. 17, pp. 7945–7957, Dec. 2014.
[11] D. Gavalas and M. Kenteris, "Evaluation of a Web recommender system in electronic and mobile tourism," *Int. J. Web Eng. Technol.*, vol. 7, no. 1, pp. 4–21, Mar. 2012.
[12] H.-N. Kim, A. Alkhaldi, A. El Saddik, and G.-S. Jo, "Collaborative user modeling with user-generated tags for social recommender systems," *Expert Syst. Appl.*, vol. 38, no. 7, pp. 8488–8496, Jul. 2011.
[13] F. Narducci, "Exploiting big data for enhanced representations in content-based recommender systems," presented at the Int. Conf. EC Web, Aug. 2013.
[14] C. Musto, F. Narducci, G. Semeraro, P. Lops, and M. de Gemmis, "A framework for explaining recommendations based on the linked open data cloud," in *Proc. 10th ACM Conf. Rec. Syst.*, 2016, pp. 151–154.
[15] K. Niemann and M. Wolpers, "Creating usage context-based object similarities to boost recommender systems in technology enhanced learning," *IEEE Trans. Learn. Technol.*, vol. 8, no. 3, pp. 274–285, Jul. 2015.
[16] M. A. Morid, M. Shajari, and A. R. Hashemi, "Defending recommender systems by influence analysis," *Inf. Retr.*, vol. 17, no. 2, pp. 137–152, Apr. 2014.
[17] Y. Blanco-Fernández, M. López-Nores, A. Gil-Solla, M. Ramos-Cabrer, and J. J. Pazos-Arias, "Exploring synergies between content-based filtering and spreading activation techniques in knowledge-based recommender systems," *Inf. Sci.*, vol. 181, no. 21, pp. 4823–4846, Nov. 2011.
[18] H. Ma, T. C. Zhou, M. R. Lyu, and I. King, "Improving recommender systems by incorporating social contextual information," *ACM Trans. Inf. Syst.*, vol. 29, no. 2, p. 9, Apr. 2011.
[19] P. Kazienko, K. Musial, and T. Kajdanowicz, "Multidimensional social network in the social recommender system," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 41, no. 4, pp. 746–759, Jul. 2011.
[20] F. Gedikli, D. Jannach, and M. Z. Ge, "How should I explain? A comparison of different explanation types for recommender systems," *Int. J. Hum.-Comput. Stud.*, vol. 72, no. 4, pp. 367–382, Apr. 2014.
[21] H. Lee and J. Kwon, "Personalized TV contents recommender system using collaborative context tagging-based user's preference prediction technique," *Int. J. Multimedia Ubiquitous Eng.*, vol. 9, no. 5, pp. 231–240, 2014.
[22] T. Hussein, T. Linder, W. Gaulke, and J. Ziegler, "Hybreed: A software framework for developing context-aware hybrid recommender systems," *User Model. User-Adapt. Interact.*, vol. 24, nos. 1–2, pp. 121–174, Feb. 2014.
[23] R. Cabacas, Y. Wang, and I.-H. Ra, "Qualitative assessment of social network-based recommender systems based on essential properties," in *Advanced Intelligent Systems*, vol. 268. Cham, Switzerland: Springer, 2014, pp. 1–11.
[24] S. K. Shinde and U. Kulkarni, "Hybrid personalized recommender system using centering-bunching based clustering algorithm," *Expert Syst. Appl.*, vol. 39, no. 1, pp. 1381–1387, Jan. 2012.
[25] A. Javari, J. Gharibshah, and M. Jalili, "Recommender systems based on collaborative filtering and resource allocation," *Social Netw. Anal. Mining*, vol. 4, no. 1, p. 234, Dec. 2014.
[26] C. He, Y. Tang, Z. Yang, K. Zheng, and G. Chen, "SRSH: A social recommender system based on Hadoop," *Int. J. Multimedia Ubiquitous Eng.*, vol. 9, no. 6, pp. 141–152, 2014.
[27] K. Wakil, R. Bakhtyar, K. Ali, and K. Alaadin, "Improving Web movie recommender system based on emotions," *Int. J. Adv. Comput. Sci. Appl.*, vol. 6, no. 2, pp. 218–226, 2015.
[28] H. Yin, B. Cui, Y. Sun, Z. Hu, and L. Chen, "LCARS: A spatial item recommender system," *ACM Trans. Inf. Syst.*, vol. 32, no. 3, p. 11, Jun. 2014.
[29] C.-L. Huang, P.-H. Yeh, C.-W. Lin, and D.-C. Wu, "Utilizing user tag-based interests in recommender systems for social resource sharing Websites," *Knowl.-Based Syst.*, vol. 56, pp. 86–96, Jan. 2014.
[30] K. McNally, M. P. O'Mahony, and B. Smyth, "A comparative study of collaboration-based reputation models for social recommender systems," *User Model. User-Adapt. Interact.*, vol. 24, no. 3, pp. 219–260, Aug. 2014.
[31] B. Li, L. Chen, X. Zhu, and C. Zhang, "Noisy but non-malicious user detection in social recommender systems," *World Wide Web*, vol. 16, nos. 5–6, pp. 677–699, Nov. 2013.
[32] A. Papadimitriou, P. Symeonidis, and Y. Manolopoulos, "A generalized taxonomy of explanations styles for traditional and social recommender systems," *Data Mining Knowl. Discovery*, vol. 24, no. 3, pp. 555–583, May 2012.
[33] M. Y. H. Al-Shamri, "Power coefficient as a similarity measure for memory-based collaborative recommender systems," *Expert Syst. Appl.*, vol. 41, no. 13, pp. 5680–5688, Oct. 2014.
[34] Y. S. Kim, "Recommender system based on product taxonomy in e-commerce sites," *J. Inf. Sci. Eng.*, vol. 29, no. 1, pp. 63–78, Jan. 2013.
[35] J. Serrano-Guerrero, E. Herrera-Viedma, J. A. Olivas, A. Cerezo, and F. P. Romero, "A Google wave-based fuzzy recommender system to disseminate information in University digital libraries 2.0," *Inf. Sci.*, vol. 181, no. 9, pp. 1503–1516, May 2011.
[36] Z. Zhang, H. Lin, K. Liu, D. Wu, G. Zhang, and J. Lu, "A hybrid fuzzy-based personalized recommender system for telecom products/services," *Inf. Sci.*, vol. 235, pp. 117–129, Jun. 2013.
[37] Z. B. Sun *et al.*, "Recommender systems based on social networks," *J. Syst. Softw.*, vol. 99, pp. 109–119, Jan. 2015.
[38] X. L. Zheng, C.-C. Chen, J.-L. Hung, W. He, F.-X. Hong, and Z. Lin, "A hybrid trust-based recommender system for online communities of practice," *IEEE Trans. Learn. Technol.*, vol. 8, no. 4, pp. 345–356, Apr. 2015.
[39] L. O. Colombo-Mendoza, R. Valencia-García, A. Rodríguez-González, G. Alor-Hernández, and J. J. Samper-Zapater, "RecomMetz: A context-aware knowledge-based mobile recommender system for movie showtimes," *Expert Syst. Appl.*, vol. 42, no. 3, pp. 1202–1222, Feb. 2015.
[40] G. M. L. Sarné, "A novel hybrid approach improving effectiveness of recommender systems," *J. Intell. Inf. Syst.*, vol. 44, no. 3, pp. 397–414, Jun. 2015.
[41] A. Tejeda-Lorente, C. Porcel, J. Bernabé-Moreno, and E. Herrera-Viedma, "REFORE: A recommender system for researchers based on bibliometrics," *Appl. Soft Comput.*, vol. 30, pp. 778–791, May 2015.
[42] R. Baraglia, P. Dazzi, M. Mordacchini, and L. Ricci, "A peer-to-peer recommender system for self-emerging user communities based on gossip overlays," *J. Comput. Syst. Sci.*, vol. 79, no. 2, pp. 291–308, Mar. 2013.
[43] F. Sanchez *et al.*, "Social and content hybrid image recommender system for mobile social networks," *Mobile Netw. Appl.*, vol. 17, no. 6, pp. 782–795, Dec. 2012.
[44] W. Hussein, R. M. Ismail, T. F. Gharib, and M. G. M. Mostafa, "A personalized recommender system based on a hybrid model," *J. Univ. Comput. Sci.*, vol. 19, no. 15, pp. 2224–2240, Sep. 2013.
[45] M. Tavakolifard and K. C. Almeroth, "Social computing: An intersection of recommender systems, trust/reputation systems, and social networks," *IEEE Netw.*, vol. 26, no. 4, pp. 53–58, Jul. 2012.
[46] B. Amini, R. Ibrahim, M. S. Othman, and M. A. Nematbakhsh, "A reference ontology for profiling scholar's background knowledge in recommender systems," *Expert Syst. Appl.*, vol. 42, no. 2, pp. 913–928, Feb. 2015.
[47] L. Quijano-Sanchez, J. A. Recio-Garcia, B. Diaz-Agudo, and G. Jimenez-Diaz, "Social factors in group recommender systems," *ACM Trans. Intell. Syst. Technol.*, vol. 4, no. 1, p. 8, Jan. 2013.
[48] P. Moradi and S. Ahmadian, "A reliability-based recommendation method to improve trust-aware recommender systems," *Expert Syst. Appl.*, vol. 42, no. 21, pp. 7386–7398, Nov. 2015.
[49] I. A. Christensen and S. Schiaffino, "Social influence in group recommender systems," *Online Inf. Rev.*, vol. 38, no. 4, pp. 524–542, 2014.
[50] J. Zhou, M. Tang, Y. Tian, A. Al-Dhelaan, M. Al-Rodhaan, and S. Lee, "Social network and tag sources based augmenting collaborative recommender system," *IEICE Trans. Inf. Syst.*, vol. E98-D, no. 4, pp. 902–910, 2015.
[51] X. Yang, Y. Guo, Y. Liu, and H. Steck, "A survey of collaborative filtering based social recommender systems," *Comput. Commun.*, vol. 41, pp. 1–10, Mar. 2014.
[52] H. Movahedian and M. R. Khayyambashi, "A tag-based recommender system using rule-based collaborative profile enrichment," *Intell. Data Anal.*, vol. 18, no. 5, pp. 953–972, Sep. 2014.
[53] J. Choi, H. J. Lee, and Y. C. Kim, "The influence of social presence on customer intention to reuse online recommender systems: The roles of personalization and product type," *Int. J. Electron. Commerce*, vol. 16, no. 1, pp. 129–154, 2011.
[54] S. M. Daneshmand, A. Javari, S. E. Abtahi, and M. Jalili, "A time-aware recommender system based on dependency network of items," *Comput. J.*, vol. 58, no. 9, pp. 1955–1966, Nov. 2015.
[55] L. Quijano-Sanchez, J. A. Recio-Garcia, and B. Diaz-Agudo, "An architecture and functional description to integrate social behaviour knowledge into group recommender systems," *Appl. Intell.*, vol. 40, no. 4, pp. 732–748, Jun. 2014.

[56] H. Zou, Z. Gong, N. Zhang, W. Zhao, and J. Guo, "TrustRank: A cold-start tolerant recommender system," *Enterprise Inf. Syst.*, vol. 9, no. 2, pp. 117–138, 2015.
[57] I. Cantador, P. Castells, and A. Bellogin, "An enhanced semantic layer for hybrid recommender systems: Application to news recommendation," *Int. J. Semantic Web Inf. Syst.*, vol. 7, no. 1, pp. 44–78, Jan. 2011.
[58] D.-H. Shih, D. C. Yen, H.-C. Lin, and M.-H. Shih, "An implementation and evaluation of recommender systems for traveling abroad," *Expert Syst. Appl.*, vol. 38, no. 12, pp. 15344–15355, Dec. 2011.
[59] G. Semeraro, P. Lops, M. De Gemmis, C. Musto, and F. Narducci, "A Folksonomy-based recommender system for personalized access to digital artworks," *J. Comput. Cultural Heritage*, vol. 5, no. 3, Oct. 2012.
[60] X. Zhou, Y. Xu, Y. Li, A. Josang, and C. Cox, "The state-of-the-art in personalized recommender systems for social networking," *Artif. Intell. Rev.*, vol. 37, no. 2, pp. 119–132, Feb. 2012.
[61] Z.-K. Zhang, T. Zhou, and Y.-C. Zhang, "Tag-aware recommender systems: A state-of-the-art survey," *J. Comput. Sci. Technol.*, vol. 26, no. 5, pp. 767–777, Sep. 2011.
[62] C. Martinez-Cruz, C. Porcel, J. Bernabé-Moreno, and E. Herrera-Viedma, "A model to represent users trust in recommender systems using ontologies and fuzzy linguistic modeling," *Inf. Sci.*, vol. 31, pp. 102–118, Aug. 2015.
[63] W. Yuan, D. Guan, S. Lee, and J. Wang, "Skeleton searching strategy for recommender searching mechanism of trust-aware recommender systems," *Comput. J.*, vol. 58, no. 9, pp. 1876–1883, Sep. 2015.
[64] M. J. A. Berry and G. S. Linoff, "Data mining methodology and best practices," in *Data Mining Techniques*, 2nd ed. Bloomington, IN, USA: Wiley, 2004, pp. 43–86.
[65] H. K. Kim, J. K. Kim, and Y. U. Ryu, "Personalized recommendation over a customer network for ubiquitous shopping," *IEEE Trans. Serv. Comput.*, vol. 2, no. 2, pp. 140–151, Apr./Jun. 2011.
[66] J. R. Quevedo, E. Montañés, J. Ranilla, and I. Díaz, "Ranked tag recommendation systems based on logistic regression," presented at the Int. Conf. Hybrid Artif. Intell. Syst., Jun. 2010. [Online]. Available: http://link.springer.com/chapter/10.1007%2F978-3-642-13769-3_29
[67] H. Wang and W. J. Li, "Relational collaborative topic regression for recommender systems," *IEEE Trans. Knowl. Data Eng.*, vol. 27, no. 5, pp. 1343–1355, May 2015.
[68] E. Karydi and K. Margaritis, "Parallel and distributed collaborative filtering: A survey," *ACM Comput. Surv.*, vol. 49, no. 2, p. 37, 2014.
[69] J. Lu, D. Wu, M. Mao, W. Wang, and G. Zhang, "Recommender system application developments: A survey," *Decision Support Syst.*, vol. 74, pp. 12–32, Jun. 2015.
[70] Y. S. Kim, "Text recommender system using user's usage patterns," *Ind. Manage. Data Syst.*, vol. 111, no. 1, pp. 282–297, 2011.
[71] Z. Zhang, H. Lin, P. Li, H. Wang, and D. Lu, "Improving semi-supervised text classification by using Wikipedia knowledge," in *Web-Age Information Management*. Berlin, Germany: Springer-Verlag, 2013, pp. 25–36.
[72] W. T. Lo, Y.-S. Chang, R.-K. Sheu, and J. Jung, "The practice of two-phase recommender system for sporting goods," *Malaysian J. Comput. Sci.*, vol. 27, no. 2, pp. 138–155, 2014.
[73] X. H. Pham, J. J. Jung, and S.-B. Park, "Exploiting social contexts for movie recommendation," *Malaysian J. Comput. Sci.*, vol. 27, no. 1, pp. 68–79, 2014.
[74] P. Levay, N. Ainsworth, R. Kettle, and A. Morgan, "Identifying evidence for public health guidance: A comparison of citation searching with Web of science and Google scholar," *Res. Synthesis Methods*, vol. 7, no. 1, pp. 34–45, Mar. 2016.
[75] P. Jacso, "Deflated, inflated and phantom citation counts," *Online Inf. Rev.*, vol. 30, no. 3, pp. 297–309, 2006.
[76] P. Jacso, "As we may search—Comparison of major features of the Web of science, scopus, and Google scholar citation-based and citation-enhanced databases," *Current Sci.*, vol. 89, no. 9, pp. 1537–1547, Nov. 2005.
[77] P. da Silva Neubert, R. S. Rodrigues, and G. R. B. Fachin, "Use of Web 2.0 tools by Ibero-American journals indexed by Web of science and scopus," *Transformation*, vol. 28, no. 2, pp. 127–141, May/Aug. 2016.
[78] S. A. Moorhead, D. E. Hazlett, L. Harrison, J. K. Carroll, A. Irwin, and C. Hoving, "A new dimension of health care: Systematic review of the uses, benefits, and limitations of social media for health communication," *J. Med. Internet Res.*, vol. 15, no. 4, p. e85, Apr. 2013, doi: 10.2196/jmir.1933.
[79] U. Hanani, B. Shapira, and P. Shoval, "Information filtering: Overview of issues, research and systems," *User Model. User-Adapt. Interact.*, vol. 11, pp. 203–259, Aug. 2001.
[80] F. M. Harper and J. A. Konstan, "The movielens datasets: History and context," *ACM Trans. Interact. Intell. Syst.*, vol. 5, p. 19, Jan. 2016.
[81] E. Çano and M. Morisio, "Characterization of public datasets for recommender systems," in *Proc. IEEE 1st Int. Forum Res. Technol. Soc. Ind. Leveraging Better Tomorrow (RTSI)*, Sep. 2015, pp. 249–257. [Online]. Available: http://ieeexplore.ieee.org/document/7325106/
[82] S. Beyer and M. Pinzger, "Synonym suggestion for tags on stack overflow," in *Proc. IEEE Conf. Program Comprehension*, May 2015, pp. 94–103. [Online]. Available: http://ieeexplore.ieee.org/document/7181436/
[83] X. Liu and K. Aberer, "SoCo: A social network aided context-aware recommender system," presented at the 22nd Int. Conf. World Wide Web WWW, May 2013, pp. 781–802.
[84] H. Liu, Z. Hu, A. Mian, H. Tian, and X. Zhu, "A new user similarity model to improve the accuracy of collaborative filtering," *Knowl.-Based Syst.*, vol. 56, pp. 156–166, Jan. 2014.
[85] J. Beel and S. Dinesh, "Real-world recommender systems for academia: The pain and gain in building, operating, and researching them," presented at the 5th Int. Workshop Bibliometric-Enhanced Inf. Retr. (BIR), Apr. 2017.
[86] D. A. Adeniyi, Z. Wei, and Y. Yongquan, "Automated Web usage data mining and recommendation system using K-nearest neighbor (KNN) classification method," *Appl. Comput. Inform.*, vol. 12, no. 1, pp. 90–108, Jan. 2016.
[87] C. Ju and C. Xu, "A new collaborative recommendation approach based on users clustering using artificial bee colony algorithm," *Sci. World J.*, vol. 2013, Oct. 2013, Art. no. 869658. [Online]. Available: https://www.hindawi.com/journals/tswj/2013/869658/cta/



ANITHA ANANDHAN received the master's degree in computer science from Bharathiar University, Coimbatore, India, and the M.Phil. degree in computer science. She is currently pursuing the Ph.D. degree with the Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur, Malaysia. She was a Lecturer with the PSG College of Technology, Coimbatore, for four years. Her research interests include online social media, data mining, and recommender system.

LIYANA SHUIB is currently a Senior Lecturer with the Department of Information Systems, University of Malaya. She has over eight years of experience as an academician and also successfully supervised Ph.D. and master's students. She is also supervising eight master's and five Ph.D. students. She has examined master's and Ph.D. theses. She has also published several articles in academic journals indexed in well-reputed databases, such as ISI-indexed and Scopus-indexed. Her research interests include information retrieval and decision support in information systems.

MAIZATUL AKMAR ISMAIL received the bachelor's degree from the University of Malaya, Malaysia, the master's degree from the University of Putra Malaysia, and the Ph.D. degree in computer science from the University of Malaya in 2010. She has over 15 years of teaching experience since she started her career in 2001 as a Lecturer with the University of Malaya. She is currently an Associate Professor with the Department of Information Systems, Faculty of Computer Science and Information Technology, University of Malaya. She was involved in various studies, leading to publication of a number of academic papers in the areas of information systems specifically on semantic Web, educational technology, and information systems. She has been actively publishing over 40 conference papers in renowned local and international conferences. A number of her papers were also published in reputable international journals. She has participated in many competitions and exhibitions to promote her research works.

GHULAM MUJTABA received the master's degree in computer science from National University, FAST, Karachi, Pakistan. He is currently pursuing the Ph.D. degree with the Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur, Malaysia. He has been an Assistant Professor with the Sukkur Institute of Business Administration, Sukkur, Pakistan, since 2006. He has vast experience in teaching and research. He was involved in well-known software house in Karachi for four years. He has also published several articles in academic journals indexed in well-reputed databases, such as ISI-indexed and Scopus-indexed. His research interests include online social networking, text mining, deep learning, and information retrieval. He is a Gold Medalist in his master's degree.