Expert Systems With Applications 103 (2018) 63–73

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

A literature review for recommender systems techniques used in microblogs

Luis Terán[a,b,*], Alvin Oti Mensah[c], Arianna Estorelli[a]

[a] University of Fribourg, Boulevard de Pérolles 90, Fribourg, Switzerland
[b] Universidad de las Fuerzas Armadas ESPE, Av. General Rumiñaui S/N, Sangolquí, Ecuador
[c] University of Bern, Neubrückstrasse 10, Bern 3012, Switzerland

ARTICLE INFO

Article history:
Received 29 October 2017
Revised 12 February 2018
Accepted 5 March 2018
Available online 6 March 2018

Keywords:
Classification
Recommender systems
Social networks
Microblogs

ABSTRACT

Online social networks (OSNs) are receiving great attention from the research community for different purposes, such as event detection, crisis management, and forecasting, among others. The increasing amount of research conducted with social networks opens the need for a classification methodology regarding trends in the field. This work does not cover all types of social networks; it focuses on the analysis of microblogs as a data source in the context of recommender systems (RSs). The main goal of this work is to provide authors with insights on the trends of academic literature reviews in the proposed context and to provide a comparison of different research approaches. The authors searched for up-to-date research papers related to RS methods using microblogs within a time period of five years, from 2012 to January 2018. Starting from 2012, a significant amount of research related to the subject field of RSs was conducted and identified by the authors of this work. After the filtering process, 39 papers were finally selected from journals and conferences in four different databases related to Internet technologies (i.e., IEEE, ACM, Science Direct, and Springer). A general classification presented in this work is then adopted and used to describe state-of-the-art social network recommendation approaches for microblogging. This work can be extended in the future to include novel methodologies and trends of RSs for microblogs.

1. Introduction

Over the last decade the Online Social Networks (OSNs) have become very popular and also they have increased both in size and in data they produce. In the context of this paper, OSNs are regarded as having two core focuses: (1) social relations and (2) user-generate d content. Despite the varying terminologies, their underlying common features are users who create their own profiles and content that can be interlinked, enabling users to connect with others, share content, and build communities.

The abundance of information created by the millions of users and their daily interaction with OSNs in which it is unimaginable for users to navigate through all of the useful information by means of search engines that are integrated within these platforms. For example, when a user types a set of keywords in a search engine, the final results are related to these keywords; users are not able to go through all results displayed, and they reduce their user experience to a small set.

Recommender systems (RSs) were developed to retrieve the top-k similar results close to users' preferences (Lu, Wu, Mao, Wang, & Zhang, 2015) by recommending the most suitable items (e.g., products or services) to target users (e.g., individuals, groups, or businesses) by examining users' interactions with items and other users and predicting users' interests. An RS can be seen as a superior form of information retrieval due to the level of personalization provided for each individual. While information filtering and retrieval provide a set of results for each individual using keywords, an RS takes into account contextual information and provides results based on individuals' tastes.

The use of data generated by social media sources is attracting researchers from the RS community to tackle different problems such as the cold start user, event recommendation, and social bookmarking systems, among others. In Carmel et al. (2009), the authors investigated personalized social searches based on users' social relations. In Kefalas, Symeonidis, and Manolopoulos (2016), the authors showed explicitly that using data from OSNs can improve recommendations. As reported in Castillejo, Almeida, and López-de Ipina (2012), the authors introduced a methodology using social relationships gathered from social networks to generate

* Corresponding author at: University of Fribourg, Boulevard de Pérolles 90, Fribourg, Switzerland.
E-mail addresses: luis.teran@unifr.ch (L. Terán), alvin.otimensah@students.unibe.ch (A.O. Mensah), arianna.estorelli@unifr.ch (A. Estorelli).
URL: http://diuf.unifr.ch/main/is/members/luis-teran (L. Terán)

https://doi.org/10.1016/j.eswa.2018.03.006
0957-4174

initial recommendations to tackle the so-called cold start user. The importance of Twitter in the social media context is not negligible. Between 2010 and 2017, the number of active members on Twitter increased by 93.3% (Statista, 2017).

In Mangal, Niyogi, and Milani (2016), the authors presented an analysis of Twitter and how it can be used in concurrent worldwide trends. There is clear evidence of worldwide coverage on the use of Twitter that concerns social entities for personalization. The relevance of Twitter in comparison to other OSNs is also important. Much attention has been given to users' profile generation and using data from Twitter to improve traditional recommendations when compared to the improvement of users' experiences through recommendations on Twitter. One example of using Twitter to enhance users' profiles was presented in Terán and Mancera (2017), in the context of political discussion.

As shown in the report of Twitter (Statista, 2017), the number of active users worldwide increased from 30 million users in the first quarter of 2010–330 million users in the third quarter of 2017. This growth goes hand in hand with the increase in the amount of research conducted using microblogging data, presented in academic literature (Kefalas, Symeonidis, & Manolopoulos, 2013).

The paper is structured as follows: The first section presents related work regarding the design of a conceptual framework for RSs. Then, the next section describes the data collection and methodology used to classify the different research papers. The fourth section shows a use case using the classification proposed for the different approaches in academic literature. Concluding remarks are presented in the final section.

2. Related work

In this section, different methodologies introduced to classify various types of RSs in the context of OSNs, emphasizing microblogging platforms, are presented. There are a handful of research works that performed the task of classifying various types of RSs. In Park, Kim, Choi, and Kim (2012) the authors presented a review and classification of different RS approaches, grouping them based on their application fields and the types of data mining techniques that were used. Contrary to works that generally classify RSs based on their application fields, this work focused on the specific field of social recommender systems (SRSs) and Twitter. Additionally, in Yang, Guo, Liu, and Steck (2014), the authors presented a survey of various types of collaborative filtering (CF) techniques for SRSs, dividing them into four categories: feedback-based, trust-based, matrix factorization-based, and nearest neighbor-based approaches.

The authors argued that most RSs focus on CF approaches and the various adaptations of these systems. This work refrains from doing so due to the potential bias introduced in the comparison of RSs for microblogging systems without properly capturing trends. Another example of classification of RSs for Twitter is presented in Kywe, Lim, and Zhu (2012). The authors introduced a taxonomy of recommendation tasks in Twitter to further describe relevant work. RSs were classified based on the entities used for recommendations, resulting in an eight-category classification framework: follower, followee, retweet, tweet, hashtag, and news. However, the study did not include a classification of existing research papers, and no provisions were made for RSs with more than one entity.

Thus, the goal of this work is to analyze the research trends on RSs used in microblogging and provide a general classification that can be extended with further methodologies and trends. This work is intended to support researchers in their analyses of literature and presents an overview of RSs in OSNs and a categorization of technologies employed in the context of microblogging.

3. Data collection and classification framework

To understand the trends on the use of RSs on microblogging, an analysis of academic literature was conducted. Additionally, a classification framework based on the results of the initial analysis was introduced.

3.1. Data collection

The selected papers were collected from four scientific databases: Science Direct, ACM Portal, IEEE Library, and Springer Link. The initial keywords used in these databases included the following: collaborative filtering, content-based filtering, hybrid-based filtering, social media, OSNs, recommender system, recommendation system, and Twitter. Depending on the results, the keywords were then permuted to find the subset that gave the best results. The final list of keywords included: RSs, recommendation system, OSNs, social media, microblogging, and Twitter.

Papers were selected if their title contained a combination of at least two keywords from the list proposed for this study. The contents of the papers were skimmed to check whether they were relevant to be selected for deeper review. During the review process, the articles selected were studied in more detail, and further unrelated papers were discarded. The final selection of papers relevant for our classification are presented in Table 1.

3.2. Classification framework

After the analysis of the selected papers, a classification framework was introduced to better understand the methodologies and trends used in the development of RSs for microblogging. Based on the analysis of the selected papers, three categories were proposed for the classification: methods used, techniques used, and recommendation types. Fig. 1 shows the classification framework to better understand the evolution of different technologies in the subject of this study.

3.3. Distribution using the classification framework

In this section, the distribution based on the three categories used in the classification framework proposed in this work is presented in Fig. 2. Fig. 2a shows the distribution of all techniques found during the analysis of research papers. It shows a clear tendency of the use of graph-based, term frequency-inverse document frequency (TF-IDF), and latent Dirichlet allocation (LDA) techniques. On the other hand, Fig. 2b shows the tendency regarding recommendation types. The tendency for studies on who to follow and tweet recommendations is presented. Finally, Fig. 2c shows the distribution of methods used along the five years that correspond to this study. It shows that most of the studies were based on content-based (CB) methods; nevertheless, the other two methods (CF and HB) cannot be neglected.

4. Use case of the classification framework

Twitter provides a function for users to be notified regarding shared information by following users they find interesting. Follow RSs try to solve the problem of searching Twitter for interesting users. Follow RSs are intended to resolve the following challenges:

- Which users to follow because their contributions are fragmented among various tiny posts.
- What influences users to follow.
- The probability that a user recommended to a target user will be accepted.
- Number of users to follow to obtain interesting information available on Twitter

Table 1
Number of papers returned from scientific databases with variation of keywords.

Library	Selection of words	Papers found	Papers selected (Initial/Final)
ACM Portal	Recommender systems, recommendation system, microblogging, Twitter	51	17/13
	Recommender systems, recommendation system, online social networks, social media, microblogging, Twitter	1	
IEEE Xplore	Recommender systems, recommendation system, online social networks, social media, microblogging, Twitter	4	18/15
	Recommender systems, recommendation system, online social networks, social media, microblogging, Twitter	26	
	Recommender systems, recommendation system, microblogging, Twitter	20	
	Recommender systems, recommendation system, online social networks, social media, microblogging	8	
Science Direct	Recommender systems, recommendation system, online social networks, social media, microblogging, Twitter	43	7/5
	Recommender systems, recommendation system, online social networks, social media, microblogging	46	
	Recommender systems, recommendation system, microblogging, Twitter	57	
Springer Link	Recommender systems, recommendation system, online social networks, social media, microblogging, Twitter	81	7/6
	Recommender systems, recommendation system, online social networks, online media	8	
	Recommender systems, recommendation system, microblogging, Twitter	55	

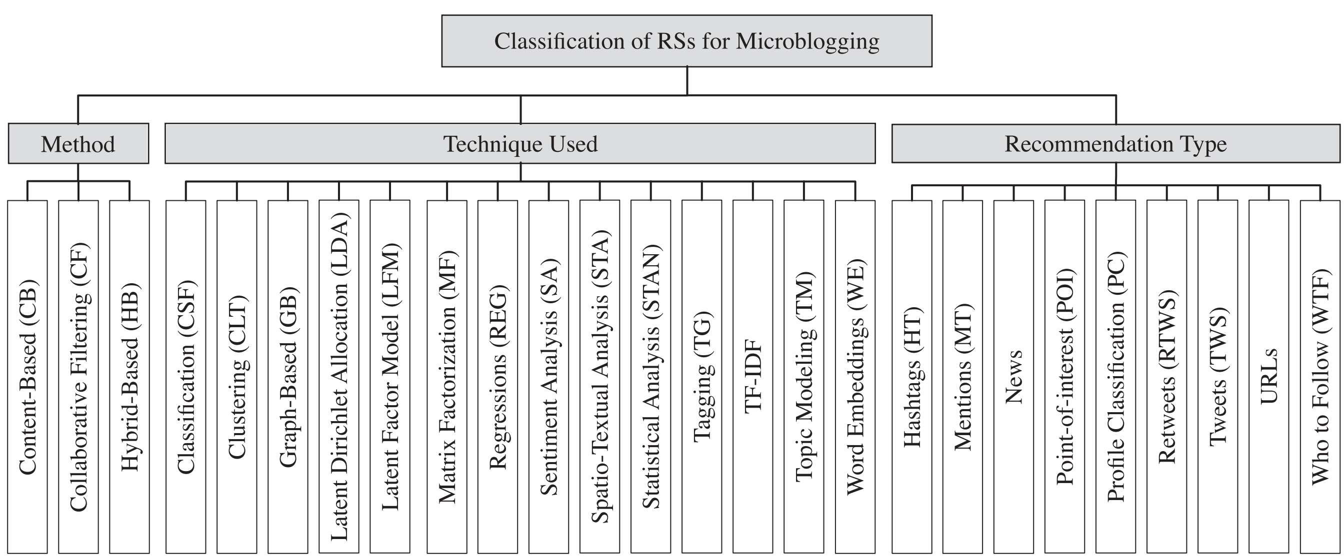


Fig. 1. Classification framework for recommender systems in microblogging.

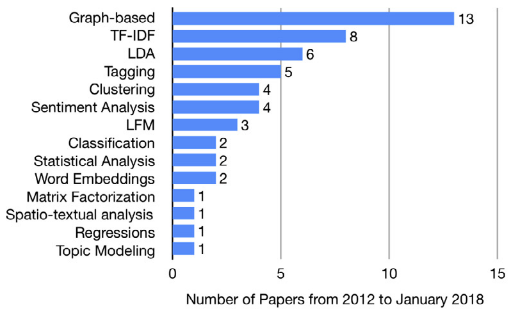
- Characteristics used to improve the accuracy of the recommendations.
- How to understand users well enough to recommend them to interesting parties.

The methods employed to solve the challenges mentioned above in the context of follow recommendation, using the classification framework presented in the previous section, are presented as follows:

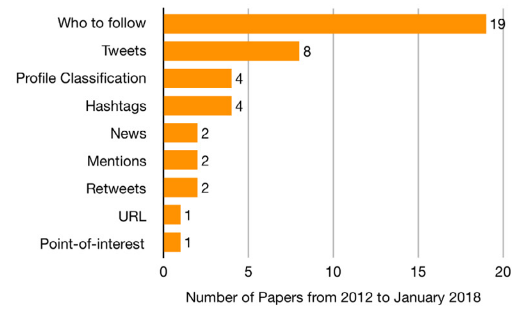
4.1. Content-based (CB) approaches

This category of follow RSs performs the task of recommending users that one could follow by finding other users with similar topics of interest (ToI) through analysis of their contents (tweets). In Celebi and Uskudarli (2012), the authors presented a CB approach for follow recommendations that sought out people with the assumption that their ToI were fragmented through their numerous tiny tweets. Thus, the tweet history of each user was processed to find keywords to be tagged and extended to find related tags, using an external tagging system. This produced a set of co-occurrences of tags (i.e., the degree of relatedness of each tag), which were indexed using the TF-IDF weighting scheme. Then, a cosine similarity was applied to find the closest users. Their social meta characteristics were computed to rank users by good blogging behaviors.

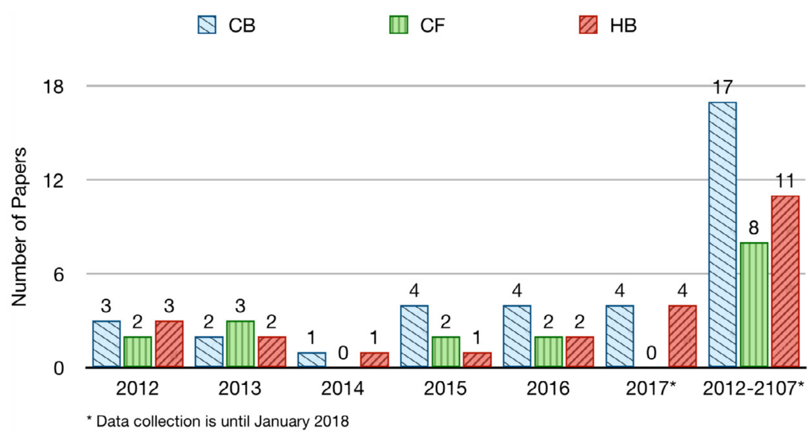
The process was carried out to produce a ranked list of recommended microbloggers given a user query. The authors argued that the follower and followee information used for recommendations in which friends of friends (FOF) were considered to be good candidates to follow was not efficient due to the large number of microbloggers users typically follow (Celebi & Uskudarli, 2012). In Yanardag Delul (2013), the authors presented another CB approach to follow RSs. The authors proposed a two-stage framework that first found potentially relevant users by extracting topics from their tweets and then took advantage of submodularity to provide rankings of recommended users, taking into account users' own topic distributions as well.



(a) Classification by Technique Used



(b) Classification by Recommendation Type



(c) Classification by Method

Fig. 2. Distribution using the classification framework.

The CB approaches are similar to each other in terms of the processes they go through to produce recommendations retrieving the topic distribution of users via their historical tweets and weighting through microblogging metrics, such as retweets, mentions, URLs, hashtags, and so on. However, they varied in the methodologies used to generate users' ToI and rankings. In Yanardag Delul (2013), the authors argued that traditional information retrieval (IR) methods such as TF-IDF were not efficient for Twitter because they assumed that the indexed documents were of a reasonable length. They also argued that topic modeling methods, such as LDA, failed to capture the behaviors of tweets (i.e., the short content of tweets did not provide sufficient word co-occurrence information for “bag of words” representations). Their approach was to extract parts of speech tags to represent the topics of tweets, in addition to hashtags, which were used as the representative topic for tweets.

The parts of speech topic model was extended to find related topics not captured from short tweets. The authors built a search engine and fed it with the complete English Wikipedia corpus, which took pre-processed tweets as inputs and returned sets of Wikipedia entities that were relevant to their respective analyzed tweets, termed “elastic search.” A freebase search was performed on an external source and combined with the related topics into super topics. This then represented the topic distribution for each user, after which a sub-modular framework was used to find related and interesting users to follow for specific users. The sub-modular framework returned a subset of users from the category that were both very similar to and diverse from the user, as much as was possible. Their ranking objective aimed to maximize the similarities between users while minimizing redundancies between them. Therefore, the most representative users would be selected, but the selected subset would also cover different aspects of the category (Yanardag Delul, 2013).

This approach presented similar and diverse users, which could mitigate the serendipity problem in RSs. In Karidi (2016), the author performed the task of topic distribution per users' tweets using knowledge graphs. Knowledge graphs are compact graphs whose nodes represent topics and edges represent relations between them. Their use profits from logical and conceptual relations between ToI. Neighboring topics in knowledge graphs are highly related. This approach took into account the true and objective relations between users' ToI. The authors stated that topic graphs provided a common basis to compare users' ToI. This scientific objective basis semantically outperformed the LDA self-topic approaches and the approach that lacked efficiency due to tweets' low word counts (Karidi, 2016).

By extracting the topics of users' historical tweets using API provided by a knowledge graph library, topic profiles were built for each user, forming topic graphs. The similarities between users inferred from the topic graphs were calculated, and the scores were ranked. Because the backbone of this approach was the knowledge graph, it should have covered all the possible ToI. The interesting feature of the CB approaches presented in this section is that they claim better, more accurate results than existing approaches of CF and hybrid RSs on Twitter.

4.2. Collaborative filtering (CF) approaches

CF is a popular recommendation method that bases its predictions and recommendations on the ratings or behavior of other similar users in the system. The fundamental assumption behind this method is that other users' opinions can be selected and aggregated in such a way as to provide a reasonable prediction of the active users' preference. Intuitively, they assume that if users agree about the quality or relevance of some items, then they will likely agree about other items (Ekstrand, Riedl, Konstan, 2011). It follows the assumption that "if users have agreed with each other in the past, they are more likely to agree with each other in the future than to agree with randomly chosen users" (Ekstrand et al., 2011).

In contrast with CB approaches, CF does not require human intervention for tagging content because item knowledge is not required. Recommendations are made based on a neighborhood whose rating profiles are most similar to that of the target user. The CF technique can be divided into user-based and item-based CF approaches:

- In the User–User CF approach, a user will receive recommendations of items liked by similar users. A user neighborhood is exploited.
- In the Item–Item CF approach, a user will receive recommendations of items that are similar to those he or she has loved in the past. An item neighborhood is exploited.

However, it is easy to see the downside of this approach because users cannot be merely described as the average of their friends. For this reason, the model-based approach takes into account the preferences of the active user as well as the neighborhood aggregate. Model-based CF has to build a mathematical model on the object, which requires deep insight into the object before prediction. Data mining and machine learning techniques are applied to find patterns from training data that can be used to make predictions for unknown ratings and uncover latent factors that explain observed ratings.

CF used in the context of follow recommendations on Twitter is based on analysis of Twitter graphs to find the users' circle (neighborhood). CF algorithms unveil links in the neighborhood that are not yet known to the target users. These approaches are solely based on the network's follow relations. In Armentano, Godoy, and Amandi (2012), the authors identified potentially interesting candidates to follow by traversing the follow relationship to a certain level. They argued that using the whole follow sub-graph provided a less personalized recommendation; thus, only close friends and relations were worthy of recommendation. The candidates were therefore a target user's followees and followers. In other words, if a user *F* follows a user who is also followed by the target user, then other people followed by *F* could be interesting to the target user.

The results of their evaluations showed that the number of common friends (links) between target users and the candidate users yielded more accurate recommendations. In Chin, Xu, and Wang (2013), the authors extended the above CF approach of using only the follow relationship subset to include features of proximity, real-world encounters, and meetings to improve explanations of the recommendations.

4.3. Hybrid-based (HB) approaches

To improve personalizing, other recommendation methods combining the benefits of two or more recommendation methodologies emerged. In Chen, Jin, and Cui (2017), the authors demonstrated this idea by extending topology-based CF with the TF-IDF content-based measure to rank the most similar users to a target user. Another example is the *TWILITE* system presented in Kim and Shim (2014). It tended to solve the problem of having insufficient information to weight the follow relationship sub-graph by taking advantage of other content information sources.

An LDA model was used to generate the topic models that represented users' ToI, and the follow relationship was decomposed using matrix factorization. The joint probability of the topic distribution with the factorization of the users in the follow matrix was then computed, giving users in the matrix with similar topic representation higher probabilities and users with similar topic models that were not in the follow matrix a probability score of zero. In Armentano, Godoy, and Amandi (2013), the authors presented which CB topic modeling features were effective before using them to profile users in the follow sub-graph. Identifying the effective features for topic modeling in combination with the topology traversing produced accurate recommendations. Their assumption was that users followed by followers of a target user's followees would potentially be interesting and should be further evaluated from a CB point of view.

Four different strategies were defined to create CB profiles of users describing the information they were likely to receive from the people to whom they subscribed: using the users' own tweets, aggregating their followers' vector representations, maintaining a vector for each followee, and grouping followers into categories using a k-means clustering algorithm. The results showed that the users' own tweets were not a good source for user profiling. In contrast, strategies that used the tweets of the followers of users for modeling interests, either individually or grouped into categories, reached high levels of precision in recommendation.

In Wu, Chen, Yu, Han, and Wu (2015), the authors used hybridization to improve trust in follow recommendations. Their approach focused on providing genuine users who were interesting to follow for cold start users. The follow relationship of the global graph was taken into consideration rather than the follow relation of the user. To remove fake accounts from contributing to the recommendation, only dense subsets of the sparse graphs were extracted to form the neighborhood. The authors argued that fake followees' accounts were themselves cold users, or followed many cold user accounts. Thus, the division into cold user and warm user (node with more interactions) subsets eliminated unreliable media nodes.

The next task was to make recommendations for the cold start users. To do this, warm user clusters were further clustered based on interests. They then clustered the cold start users based on the warm users' cluster subsets. More specifically, for each cold start user node, a content vector was extracted from the user's tweets. The content vector of each warm cluster was extracted in the same way.

A cold start user node would be assigned to the warm-start cluster with the most similar content vector. Heterogeneity came into play when the similarity score to each cluster was computed, which represented different ToI strengths of each user to different clusters, depicting that users had varying degrees of agreeability to various ToI within different groups. Similarity within the cluster to the user was then ranked to give the recommendation. In Wu et al. (2015), the authors used the implicit feedback from trusted neighborhoods (actions in the follow relationship sub-graph) to make recommendations based on how much the user implicitly trusted the candidate users based on content of tweets, retweets, and so on.

Another hybrid recommender was presented in Yamamoto, Kumamoto, and Nadamoto (2015). The proposed system presented the argument that using only extracted ToI from users' tweets without knowing the sentiment of the user led to erroneous assumptions because at times, users advertised negative sentiments on certain topics. Thus, the RS sought to recommend similar users from the neighborhood with similar sentiments, using an eight-dimensional sentiment analysis as opposed to a positive and negative sentiment analysis. Users had various feelings about different topics, and a polarized dimensionality was insufficient.

4.3.1. Feature-based approaches

These types of RSs restructured the task of recommending from finding friends of interest to determining the features that influence the acceptance of recommendations or finding features available in the social network that are most effective to be used for recommending. The effective features identified were thus used for the task of recommendation. In Zhang, Fang, Chen, and Tang (2015), the authors investigated how new following links formed in social networks triggered the formation of other neighboring links.

The assumption was that when a user *A* followed another user *C*, this created a probability for *A′s* follower *B*, to discover *C*. *A*, *B*, and *C* form a basic triadic structure. The authors showed the link correlations in five different triadic structures. Their approach was that given the preexisting link between *A* and *B*, and the new link from *A* to *C* added at time t', they found the ratio of a new link from *B* to *C*, created within time frame *d* after t' for each triadic structure, where *d* is a time delay parameter empirically set at seven days. Consequently, they showed how the new follow link diffused through the network.

This approach proved the effectiveness of the follow feature of microblogs in recommendations used for FOF algorithms. Twenty-four other triadic structures were studied to form the conclusion that in the context of followee diffusion, a relationship directed from *A* to *C* improved the likelihood that *B* would follow *C* triggered by *B* following *A*. A maximum likelihood algorithm was developed to maximize the likelihood of the observed pattern and predict future links (followees) for users.

In Man, Shen, and Cheng (2012), the authors focused on finding features that drive users to accept celebrity recommendations on microblogging sites. Three factors in user and celebrity interactions were considered: (1) popularity, (2) structure similarity, and (3) topic similarity. The number of followers of one celebrity defined the popularity, and the structure similarity was computed using the follow relation sub-graph. For topic similarity, LDA was used to get the ToI of the users and celebrities.

The results showed that popularity and structure similarity were two major factors for users' acceptance, and topic similarity was less significant when users chose to follow celebrities in the recommended list. This was more generalized in Wu, Sorathia, and Prasanna (2012a), where the authors examined the most influential factors that users might consider in selecting followees (not only celebrities). After features were found, they were used to recommend items that matched the users' preferences.

The authors generated five recommendation engines providing ranked lists of users using different features, including item category, item popularity, follows, semantic keywords, and action influences. The results showed that item popularity was again the most prominent feature in users' adoption of recommendations. This reflects the celebrity effect prevalent on microblogging sites. Additionally, the results indicated that retweets, mentions, and comments had positive influences on the following activity of one's followers to one's followees.

4.4. Recommendation types

In the context of microblogging, and in particular on the social network Twitter, various forms of actions among users are defined. These actions produce different kinds of content, rightfully referred to as "user-generated content" (UGC), which proliferates the network. UGC includes tweets, retweets, hashtags, mentions (comments or replies), follows, URLs, and news. These items and related ones make up social entities. In this section, various technologies used for the recommendation of these social entities on Twitter are presented.

4.4.1. Tweet recommendations

After finding a set of interesting users to follow, all tweets from followees are displayed on the user's homepage, referred to as the timeline, in chronological order. Users might miss interesting information if one user posts many tweets in a row. RSs emerged on Twitter to solve this problem by recommending the most interesting tweets to users, or the most interesting from each followee based on the users' preferences.

Other types of RSs seek to encourage users to engage more in the network by recommending tweets to which users might reply. In Yuan, Huang, Sun, Li, and Xu (2015), the authors proposed a CF approach to tweet recommendations for cold start users, which models users' interest in tweets as changing over time along with the popularity of tweets' topics. As tweets age, their popularity gradually dies out. When a cold start user is identified, the members of his or her community are used to estimate topic interests in the form of tags. Then, tweets that correspond to the current ToI of the user's community are promoted. In Sudo, Nagasaka, Kobayashi, Taniguchi, and Takano (2013), the authors aimed to recommend tweets that encouraged interaction on the network. The assumption was that similar topics encouraged target users to reply to tweets. The degree of tweet similarity between a source and target user was estimated.

The familiarity of source and target users from community structure was also estimated. With the notion that users rarely replied to messages sent by strangers, and this being the same with replies to tweets by unknown users, target users would easily reply to users known to them. Tweets from the followees sub-graph with higher joint estimates of tweet similarity and user familiarity from the community were promoted. In Sun and Zhu (2013), the authors focused on recommending more personalized tweets. They argued that users followed others because they were interested in aspects of their tweets, and these aspects were checked against the current tweets of followees. Tweets with such aspects matching were promoted. They achieved dynamism by utilizing ego networks, which is a sub-graph of only direct neighbors.

In Alawad, Anagnostopoulos, Leonardi, Mele, and Silvestri (2016), the authors presented the idea of recommending tweets that were not visible but could be very interesting to users. This could happen when no one in the user's circle linked to a tweet. They achieved this by extending the user's ego network to a depth of two (that is, a neighborhood two hops from the user was constructed). Contents were analyzed, and structural similarity scores were calculated for recommendation. In Ma, Jia, Zhang, and Lin (2017), the authors argued that users' interests were not simply the aggregated interests of their communities or follow relations. Users utilize tags to annotate themselves in their profiles, which should be taken into consideration when finding interesting tweets. In cases where users' annotated tags did not exist, the authors proposed the use of TF-IDF to identify the tags to annotate users. Tweets in the follow relation neighborhood of users with higher similarity scores for the annotated tags were identified for recommendation.

In addition to using tweets and retweets and exploring the follow relation neighborhood for latent features, the most recent paper on Twitter (Cui, Du, Shen, Zhou, & Li, 2017) that was found for this study used a bipolar sentiment classifier to recommend interesting tweets of similar users with equal sentiments on specific topics. However, as already pointed out in Yamamoto et al. (2015), a polarized sentimental dimensionality was not effective in capturing users' behavior.

4.4.2. Retweet recommendations

After being presented with tweets from followees, users can choose to share tweets they find interesting in their ego networks. Some Twitter accounts that serve as information sources (portal accounts) do not provide tweets of their own but rather forward tweets of other users that their followers find interesting. They act as bridges between their followers and the diverse information sources they do not know. It is important for such portal accounts to know which tweets or topics are considered interesting in their network. Retweet RSs are available to help users on Twitter to know which groups of tweets are better suited to be shared in their communities. In this context, Zhao and Tajima (2014) investigated which metrics keep users interested in retweets.

The aim was to make recommendations for portal accounts, tweets of interest for their followers, and an appropriate manner in which to share these tweets. The authors evaluated four approaches to determine what retweeting behavior was desired by the followers after finding interesting tweets. Too few retweets makes the source not interesting, and too many retweets floods the users' timelines. Interesting topics were found by using TF-IDF on all followers' vectors, and the representative ToI of the followers were determined. The basic strategies of their algorithms to determine the appropriate retweeting interval and tweet number were: (1) timeliness-oriented (online) approaches, which retweeted immediately after getting tweets, and (2) selection-quality-oriented (near-online) approaches, which prioritized selection quality.

The evaluation was calculated based on the ratio of tweets shared by the algorithm to the ratio of tweets shared by the friends of a user within a given time interval. The authors concluded that quality-oriented algorithms outperformed online algorithms. They only achieved higher selection quality with acceptable delays. However, delays could affect the interest in the tweet to followers who may see the tweet from other sources.

4.4.3. Hashtag recommendations

In microblogging, hashtags are topic words placed after a "#" symbol to highlight social events or hot topics. Hashtags highlight the topic of tweets and make tweets easily searchable and understood by others. In Song, Meng, and Zheng (2015), the authors proposed an RS to help users select appropriate hashtags that had the tendency to diffuse more rapidly. In addition to the semantic similarity between the hashtag and its corresponding tweets, the hashtag's user acceptance degree and development tendency were taken into consideration. The item count of the related existing hashtags were used for the ranking.

In Tajbakhsh and Bagherzadeh (2016), the authors dealt with the problem of users using various forms of hashtags to represent the same concept. An example presented was #friends and #frnds used to represent *Friends*, the famous American TV comedy series, by two different users. Different from traditional TF-IDF, the authors suggested adding semantic meaning to the hashtags from the tweets they represented to form a semantic TF-IDF vector that could be used along with any similarity measure. Terms that were scored zero by traditional TF-IDF were further investigated for semantic similarity, which may have been due to shortness of words. The two were used to weight the similarity between tweets and thus the relatedness of the hashtags used to represent them. The interesting factor of this approach was that it proved to be six times more efficient in finding the relatedness of hashtags to current tweets than traditional TF-IDF models.

A novel work, (Gong, Zhang, Han, & Huang, 2017), deviated from the normal "bag of words" similarity of tweets to tags. The authors pointed out the differences in meaning of a phrase from a set of words. The main problem was that they observed that aligning words with hashtags frequently resulted in the loss of the inherent meaning that microblog users wished to express. Thus, the tags chosen represented the phrase and not the words. They used phrase-based topic modeling for tags to recommend appropriate hashtags. Regarding phrases as units could enhance the topic modeling performance of tweets and user profiles.

4.4.4. Mention recommendations

In Twitter-like social networking services, people can use the @ symbol to mention other users in tweets. Once the @ symbol is put in front of a username, an alert will be sent to that user signaling that a microblogger is commenting directly to him or her. In Gong, Zhang, Sun, and Huang (2015), the authors made recommendations for mentions on Twitter based on this review. They argued that with millions of users on Twitter, there should be an automatic user suggestion function based on incoming messages that users intend to mention. They explained their approach as follows: Given a microblog and its author, the @ recommendation task would discover a list of candidate users to complete the tweet. First, a generative model was used to learn the joint distribution of the topics, the microblogs, and the mentioned users. Then, the learned probability was used to generate candidate lists.

They then investigated which LDA modeling would best represent the ToI to give better recommendations of people to mention. Three different variants of LDA algorithms were implemented and tested. The At Topic Translation Model (A-TTM) assumed that each microblog contained a mixture of topics. The At User Topic Translation Model (A-UTTM) assumed that users tweeted about a mixture of topics, represented by a topic distribution, and each microblog had a single topic label. Finally, the At User-User Topic Translation Model (A-UUTTM) assumed that when users posted a microblog, they first generated the words in the microblog.

After the selection of a user to mention, based not only on the topic and topic words in the microblog but also on the microblogs of the mentioned user, their results showed that all three LDA topic modeling variants outperformed four existing state-of-the-art modeling approaches: Link-pLSA-LDAR (Nallapati & Cohen, 2008), frequency descending (which recommends users who are frequently mentioned by the author), citation translation mode (Huang et al., 2012), and ranking (Wang et al., 2013). A-UUTTM achieved the highest accuracy score for the recommended users selected from the list. This approach, however, assumed that the mention was made at the end of the tweet.

4.4.5. URL recommendations

Good tweeting behavior can be categorized as when URLs are included within tweets, providing more information on short tweets. This metric was used by some recommendation algorithms to weigh and rank users. URLs included in tweets should, however, relate to the content. The link recommender from Yazdanfar and Thomo (2013) sought to recommend URLs to Twitter users. It depended on tweets by users that contained URLs and used the hashtags of tweets to represent the tweets' topics. They used the entire Twittersphere's tweets with URLs and hashtags as their neighborhood. The similarity was computed between users whose tweets contained the same URLs and their hashtag representations, as well as the hashtags of the corresponding tweets of users. Excluding tweets with wrong URLs and hashtags could potentially improve the performance, in addition to weighting users' similarities.

4.4.6. News recommendations

In Nagaki, Yamaguchi, Amagasa, and Kitagawa (2016), the authors proposed improvements of news recommendations by classifying them as local or global and tailoring news to users in the local geography of the news areas, whereas all users are interested in global news. In Natarajan and Moh (2016), the authors proposed personalizing news recommendations by building a hybrid user profile that involved the analysis of clickthroughs, users' tweets, and users' friends, incorporating the importance of temporal dynamics with location preferences. Users were allowed to choose the ratio of popular news against trending news they desired. These preferences with locations of interest were taken into consideration in the recommendations.

4.4.7. Hybrid entity recommendations

This section presents a description of RSs that differs from the classic approaches. Classic models use one algorithm to recommend one entity, and hybrid approaches use the same algorithm to recommend more than one entity at a time. In Xiao, Du, Zhu, and Li (2012), the authors proposed an RS algorithm that made recommendations to users regarding follows and tweets. This was achieved by mapping how strong users were connected to the various tags they chose to annotate themselves. Upon completion of this strength association, a so-called tag map, recommendations could be made by proposing users with similar strengths to tags or their tweets.

For new users, the top-n users strongly connected to the tags they chose to represent themselves were recommended. In Karidi, Stavrakas, and Vassiliou (2017), the authors presented a recommendation technique that used knowledge graphs to semantically construct user profiles from tweets and from this user profiling recommended followees and tweets. A user profiling unit used a knowledge graph to find ToI to form a user profile vector. A similarity score was calculated over these user profile vectors by measuring the sub-graph overlap of the ToI in the knowledge graph. Tweets were also represented by their topics, and using the same knowledge graph, tweets that were related to the user profile vector were recommended. The backbone of the follow and tweets RSs was the construction of a user profile on which both the selection of users and tweets were based.

In Wu, Gong, Rand, and Raschid (2012b), the authors used link prediction to make recommendations for retweets and mentions. The idea for this entity hybridization was to find the top-n users who were largely influenced by users, and would retweet or mention those users, then recommending tweets and mentions of the focal user (the influential user) to these users.

The classifications presented in this section are summarized in Table 2. To give an indication of the trends of RSs on Twitter, the selected papers are presented in descending order, from 2018 to 2012.

4.5. Discussion

The results presented in Fig. 2 and Table 2 show the tendencies regarding the use of RSs in microblogs. Given the nature of microblogs, which is a combination of blogging platforms and instant messaging that allows users to post short messages to be shared with an audience online, the use of CB approaches is mostly used. The results suggest a tendency to use of graph-based, TF-IDF, LDA, tagging, clustering, sentiment analysis, and LMF techniques. This could be also explained given the two major groups of recommendation types presented in this study, which are: who to follow and tweets recommendations. The first type is an approach to tackle one of the most influential factors that a user might consider in selecting followees. In this scenario, the goal is to recommend other users that match the active user preferences. In this case, graph-based techniques are mainly used. The second major group of recommendation type is related to content (tweets). The goal in this case is to recommend relevant content to an active user according to his/her specific behaviour and preferences. In this case, TF-IDF techniques are the most used.

During the design of microblog platforms, additional data is encoded using open-standard file formats, such as JavaScript Object Notation (JSON). As an example, Twitter platform developed a number of object definitions, including: tweet object, user object, entities object, extended entities object, and geo-objects. Each object type has a set of attributes (e.g., name, location, coordinates, source, media). The attributes can be used for different purposes and recommendation types, like the ones identified in this study, which are profiles classification, hashtags, news, mentions, retweets, URLs, and points-of-interest (POI). An example of this recommendation type is presented in Kefalas and Manolopoulos (2017). The authors propose two unified models to provide reviews and POI recommendations. The approach considers the spatial, textual and temporal factors simultaneously.

Another example of a different recommendation type is presented in Terán and Mancera (2017). The authors use data from Twitter platform to generate dynamic profiles of political actors to develop the so-called voting advice applications (VAAs). The authors use sentiment analysis to characterize the profiles of candidates. The application provides voters relevant information on candidates and political parties by comparing their political interests with parties or candidates.

Similar users' features in microblogs are dominant through all the other entities (i.e., tweet, retweets, and mentions). These features go on to boost the effects its gets from user's interactions. Thus, the ToI distribution between users in microblogs is a powerful feature and can be seen as type of closed feedback signal. Finding methods that accurately models this distribution will increase the system explain ability as well. Additionally, it pushes the two major entities presented in this study (tweets and WTF). They can be used also in different domains, such as topic trending, forecasting, location-based, and voting advice applications, among others.

5. Conclusions

This paper presents a literature review of the state of the art on recommendations provided within microblogging. This systematic review was prompted given the development of Twitter as one of the most popular microblogging OSNs and its high usage by the public, in which researchers have shown their interest in using their data source for different purposes.

Twitter's limit of 140 characters (in late 2017, it was extended to 280 characters) makes it preferable for analysis of opinions because users tend to choose words that will convey their interests easily. On the other hand, the limited length of posts also presents challenges for RSs. Users struggle to gain more space for their tweets by using short forms of words that are not grammatically correct or jargon that only friends can easily understand. These discrepancies and other various challenges, such as cold start users, sparsity, and trust, among others, are to be examined.

Various RS approaches used from 2012 to January 2018 were reviewed to identify the latest techniques employed to improve accuracy, performance, and personalization for the different recommendation types performed. To achieve this, a classification framework was introduced to categorize various approaches and methodologies available, as well as to identify trends in research. This framework was then used on a list of 39 research papers. Section 4 describes in detail the use of this classification over the research papers used for this work. The results are summarized in Table 2.

From these classifications, it was possible to identify the distribution of all techniques found during the analysis of research papers. The most used techniques are graph-based, TF-IDF, and LDA. On the other hand, another tendency regarding recommendation types shows that studies on who to follow and tweet recommendations are attracting the interest of the RS community.

Other works turned to semantics representation using knowledge graphs and sentiment analysis, which evolved from bipolar to multi-dimensional sentiment analysis. Topic modeling evolved from using "bags of words" to early phase suggestions of representing the ToI with phrase-based topic modeling. There are branches of work dedicated to understanding user behavior in accepting recommended entities. They perform evaluations using various metrics to understand what influences users to accept particular entities. The development of RSs for Twitter should take

Table 2
Selected papers from Scientific databases using variation of keywords.

		Technique used														Recomm. type									Data used										
Paper & year	Methods	GB	TF-IDF	LDA	Tagging	CLT	SA	LFM	CSF	STAN	WE	MF	STA	REG[c]	TM	WTF	TWTS	PC	HT	News	MT	RTWS	URLs	POI	Users	TWS	RTWS	HT	PM[a]	Location	Followees	Followers	TS	BH[b]	URLs
Gong et al. (2017)	CB			x															x						x	x		x							
Dey, Shrivastava, Kaushik, and Subramaniam (2017)											x								x						x	x		x							
Karidi et al. (2017)		x														x	x								x	x			x						
Recalde et al. (2017)											x					x									x	x			x						
Karidi (2016)		x														x									x	x									
Tajbakhsh and Bagherzadeh (2016)		x	x																x						x	x		x							
Nagaki et al. (2016)		x	x																	x					x	x				x					
Natarajan and Moh (2016)			x		x										x					x					x	x				x	x			x	
Gutierrez and Poblete (2015)						x	x											x							x	x									
Yamamoto et al. (2015)						x	x									x									x	x					x				
Song et al. (2015)				x										x					x						x	x		x			x	x			
Gong et al. (2015)				x																	x				x	x							x		
Zhao and Tajima (2014)			x																			x			x	x	x								
Subercaze, Gravier, and Laforest (2013)		x																x							x	x									
Yanardag Delul (2013)		x			x											x									x	x	x								
Xiao et al. (2012)		x			x											x	x								x	x			x						
Celebi and Uskudarli (2012)			x		x											x									x	x	x						x		x
El-Arini, Paquet, Herbrich, Van Gael, and Agüera y Arcas (2012)															x			x							x		x	x	x						
Kefalas and Manolopoulos (2017)	CF												x											x	x	x				x					
Chen, Cui, and Jin (2016)								x								x									x	x	x				x		x		
Zhang et al. (2015)		x																																	
Yuan et al. (2015)												x					x								x						x	x			
Chin et al. (2013)		x														x									x						x	x			
Sun and Zhu (2013)		x															x								x		x		x		x				
Yazdanfar and Thomo (2013)						x																	x		x	x		x							x
Wu et al. (2012a)										x						x									x	x			x		x	x			
Armentano et al. (2012)		x														x									x						x	x			
Eliacik and Erdogan (2018)	HB	x					x									x									x	x					x	x			
Cui et al. (2017)							x	x									x								x	x	x				x	x			
Ma et al. (2017)			x									x					x								x	x			x		x	x			
Chen et al. (2017)		x	x													x									x	x					x	x			
Takemura and Tajima (2016)									x		x							x							x				x		x				
Alawad et al. (2016)		x															x								x	x					x	x			
Wu et al. (2015)		x	x			x										x									x	x					x	x			
Kim and Shim (2014)				x								x				x									x	x					x	x			
Sudo et al. (2013)		x		x													x								x	x	x				x	x	x		
Armentano et al. (2013)		x	x			x										x									x	x					x	x			
Wu et al. (2012b)		x								x											x	x			x	x	x				x		x		
Man et al. (2012)										x						x									x	x			x		x	x			

[a] PM: Profile Metadata
[b] BH: Browser History

this into account to improve recommendations or perform further evaluation to confirm these claims.

It was also observed that recent works regarding RSs used CB and hybrid approaches. CB approaches relate to improving topic modeling algorithms to accurately infer users' interests, while other authors argued that users' self-descriptions were critical to user profiling. The distribution of methods used along the five years that correspond to this study shows a concentration of research related to CB methods; nevertheless, the other two methods (CF and HB) cannot be neglected.

The evaluation of these recommendation approaches should be the research direction itself because some researchers contradict each other in their evaluation methods. CF supplemented simple FOF approaches (because they do not provide enough information about users) with state-of-the-art CB approaches, which has increased the number of hybrid systems and decreased the number of CF-only RSs. A new perspective of CF approaches is understanding the strength and influence of users, identifying users to follow, and including trust in RSs for communities.

Various approaches with CF also seek to solve the cold start user through community structure to predict users' interests. Research has been done to evaluate the behavior of users in a community if the community can provide more information than just the user.

Finally, this work intended to contribute by helping researchers through the use of the proposed classification framework, providing direction in finding relevant work on the field; nevertheless, it can be extended including up-to-date research papers and adding other databases that publish research work in the field of RSs in microblogs as the work presented in Liu, Yu, Wei, and Ning (2018).

Acknowledgment

The authors of this work would like to thank the Information System Research Group (http://diuf.unifr.ch/is) at the University of Fribourg for their valuable contributions.

References

Alawad, N. A., Anagnostopoulos, A., Leonardi, S., Mele, I., & Silvestri, F. (2016). Network-aware recommendations of novel tweets. In *Proceedings of the thirty-ninth international ACM SIGIR conference on research and development in information retrieval* (pp. 913–916). ACM.
Armentano, M. G., Godoy, D., & Amandi, A. (2012). Topology-based recommendation of users in micro-blogging communities. *Journal of Computer Science and Technology, 27*(3), 624–634.
Armentano, M. G., Godoy, D., & Amandi, A. A. (2013). Followee recommendation based on text analysis of micro-blogging activity. *Information Systems, 38*(8), 1116–1127.
Carmel, D., Zwerdling, N., Guy, I., Ofek-Koifman, S., Har'El, N., Ronen, I., et al. (2009). Personalized social search based on the user's social network. In *Proceedings of the eighteenth ACM conference on information and knowledge management* (pp. 1227–1236). ACM.
Castillejo, E., Almeida, A., & López-de Ipina, D. (2012). Social network analysis applied to recommendation systems: Alleviating the cold-user problem. In *Proceedings of the international conference on ubiquitous computing and ambient intelligence* (pp. 306–313). Springer.
Celebi, H. B., & Uskudarli, S. (2012). Content based microblogger recommendation. In *Proceedings of the international conference on privacy, security, risk and trust (PASSAT), and international confernece on social computing (Socialcom)* (pp. 605–610). IEEE.
Chen, H., Cui, X., & Jin, H. (2016). Top-k followee recommendation over microblogging systems by exploiting diverse information sources. *Future Generation Computer Systems, 55*, 534–543.
Chen, H., Jin, H., & Cui, X. (2017). Hybrid followee recommendation in microblogging systems. *Science China Information Sciences, 60*(1), 012102.
Chin, A., Xu, B., & Wang, H. (2013). Who should i add as a friend?: A study of friend recommendations using proximity and homophily. In *Proceedings of the fourth international workshop on modeling social media* (p. 7). ACM.
Cui, W., Du, Y., Shen, Z., Zhou, Y., & Li, J. (2017). Personalized microblog recommendation using sentimental features. In *Proceedings of the IEEE international conference on big data and smart computing (BIGCOMP)* (pp. 455–456). IEEE.
Dey, K., Shrivastava, R., Kaushik, S., & Subramaniam, L. V. (2017). Emtagger: a word embedding based novel method for hashtag recommendation on twitter. *IEEE International Conference on Data Mining (ICDM) ACUMEN Workshop*. IEEE.
Ekstrand, M. D., Riedl, J. T., & Konstan, J. A. (2011). Collaborative filtering recommender systems. *Foundations and Trends® in Human–Computer Interaction, 4*(2), 81–173.
El-Arini, K., Paquet, U., Herbrich, R., Van Gael, J., & Agüera y Arcas, B. (2012). Transparent user models for personalization. In *Proceedings of the eighteenth ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 678–686). ACM.
Eliacik, A. B., & Erdogan, N. (2018). Influential user weighted sentiment analysis on topic based microblogging community. *Expert Systems with Applications, 92*, 403–418.
Gong, Y., Zhang, Q., Han, X., & Huang, X. (2017). Phrase-based hashtag recommendation for microblog posts. *Science China Information Sciences, 60*(1), 012109.
Gong, Y., Zhang, Q., Sun, X., & Huang, X. (2015). Who will you@? In *Proceedings of the twenty-fourth ACM international on conference on information and knowledge management* (pp. 533–542). ACM.
Gutierrez, F. J., & Poblete, B. (2015). Sentiment-based user profiles in microblogging platforms. In *Proceedings of the twenty-sixth ACM conference on hypertext & social media* (pp. 23–32). ACM.
Huang, W., Kataria, S., Caragea, C., Mitra, P., Giles, C. L., & Rokach, L. (2012). Recommending citations: translating papers into references. In *Proceedings of the twenty-first ACM international conference on information and knowledge management* (pp. 1910–1914). ACM.
Karidi, D. P. (2016). From user graph to topics graph: Towards twitter followee recommendation based on knowledge graphs. In *Proceedings of the IEEE thirty-second international conference on data engineering workshops (ICDEW)* (pp. 121–123). IEEE.
Karidi, D. P., Stavrakas, Y., & Vassiliou, Y. (2017). Tweet and followee personalized recommendations based on knowledge graphs. *Journal of Ambient Intelligence and Humanized Computing, 8*, 1–15.
Kefalas, P., & Manolopoulos, Y. (2017). A time-aware spatio-textual recommender system. *Expert Systems with Applications, 78*, 396–406.
Kefalas, P., Symeonidis, P., & Manolopoulos, Y. (2013). New perspectives for recommendations in location-based social networks: Time, privacy and explainability. In *Proceedings of the fifth international conference on management of emergent digital ecosystems* (pp. 1–8). ACM.
Kefalas, P., Symeonidis, P., & Manolopoulos, Y. (2016). A graph-based taxonomy of recommendation algorithms and systems in LBSNs. *IEEE Transactions on Knowledge and Data Engineering, 28*(3), 604–622.
Kim, Y., & Shim, K. (2014). TWILITE: A recommendation system for twitter using a probabilistic model based on latent Dirichlet allocation. *Information Systems, 42*, 59–77.
Kywe, S. M., Lim, E.-P., & Zhu, F. (2012). A survey of recommender systems in twitter. In *Proceedings of the international conference on social informatics* (pp. 420–433). Springer.
Liu, L., Yu, S., Wei, X., & Ning, Z. (2018). An improved apriori-based algorithm for friends recommendation in microblog. *International Journal of Communication Systems, 31*(2).
Lu, J., Wu, D., Mao, M., Wang, W., & Zhang, G. (2015). Recommender system application developments: A survey. *Decision Support Systems, 74*, 12–32.
Ma, H., Jia, M., Zhang, D., & Lin, X. (2017). Combining tag correlation and user social relation for microblog recommendation. *Information Sciences, 385*, 325–337.
Man, T., Shen, H.-W., & Cheng, X.-Q. (2012). The untold story behind the recommendation in micro-blogging network. In *Proceedings of the second international conference on cloud and green computing (CGC)* (pp. 760–764). IEEE.
Mangal, N., Niyogi, R., & Milani, A. (2016). Analysis of users' interest based on tweets. In *Proceedings of the international conference on computational science and its applications* (pp. 12–23). Springer.
Nagaki, S., Yamaguchi, Y., Amagasa, T., & Kitagawa, H. (2016). Local attention analysis and prediction of online news articles in twitter. In *Proceedings of the adjunct proceedings of the thirteenth international conference on mobile and ubiquitous systems: Computing networking and services* (pp. 136–141). ACM.
Nallapati, R., & Cohen, W. W. (2008). Link-PLSA-LDA: A new unsupervised model for topics and influence of blogs. In *Proceedings of the international AAAI conference on web and social media* (pp. 84–92). Association for the Advancement of Artificial Intelligence.
Natarajan, S., & Moh, M. (2016). Recommending news based on hybrid user profile, popularity, trends, and location. In *Proceedings of the international conference on collaboration technologies and systems (CTS)* (pp. 204–211). IEEE.
Park, D. H., Kim, H. K., Choi, I. Y., & Kim, J. K. (2012). A literature review and classification of recommender systems research. *Expert Systems with Applications, 39*(11), 10059–10072.
Recalde, L., Mendieta, J., Boratto, L., Teran, L., Vaca, C., & Baquerizo, G. (2017). Who you should not follow: Extracting word embeddings from tweets to identify groups of interest and hijackers in demonstrations. *IEEE Transactions on Emerging Topics in Computing, PP*(99), 1–15.
Song, S., Meng, Y., & Zheng, Z. (2015). Recommending hashtags to forthcoming tweets in microblogging. In *Proceedings of the IEEE international conference on systems, man, and cybernetics (SMC)* (pp. 1998–2003). IEEE.
Statista (2017). Number of monthly active Twitter users worldwide from 1st quarter 2010 to 3rd quarter 2017 (in millions). Retrieved from https://www.statista.com/statistics/282087/numberof-monthly-active-twitter-users/.
Subercaze, J., Gravier, C., & Laforest, F. (2013). Towards an expressive and scalable Twitter's users profiles. In *Proceedings of the IEEE/WIC/ACM international joint conferences on web intelligence (WI) and intelligent agent technologies (IAT): 1* (pp. 101–108). IEEE.

Sudo, K., Nagasaka, S., Kobayashi, K., Taniguchi, T., & Takano, T. (2013). Encouraging user interaction of social network through tweet recommendation using community structure. In *Proceedings of the conference on technologies and applications of artificial intelligence (TAAI)* (pp. 300–305). IEEE.

Sun, J., & Zhu, Y. (2013). Microblogging personalized recommendation based on ego networks. In *Proceedings of the IEEE/WIC/ACM international joint conferences on web intelligence (WI) and intelligent agent technologies (IAT): 1* (pp. 165–170). IEEE Computer Society.

Tajbakhsh, M. S., & Bagherzadeh, J. (2016). Microblogging hash tag recommendation system based on semantic TF-IDF: Twitter use case. In *Proceedings of the IEEE international conference on future internet of things and cloud workshops (FiCLOUDW)* (pp. 252–257). IEEE.

Takemura, H., & Tajima, K. (2016). Classification of twitter accounts into targeting accounts and non-targeting accounts. In *Proceedings of the twenty-seventh ACM conference on hypertext and social media* (pp. 291–296). ACM.

Terán, L., & Mancera, J. (2017). Dynamic profiles using sentiment analysis for VAA's recommendation design. *Procedia Computer Science, 108*, 384–393.

Wang, B., Wang, C., Bu, J., Chen, C., Zhang, W. V., Cai, D., & He, X. (2013). Whom to mention: Expand the diffusion of tweets by@ recommendation on micro-blogging systems. In *Proceedings of the twenty-second international conference on world wide web* (pp. 1331–1340). ACM.

Wu, H., Sorathia, V., & Prasanna, V. K. (2012a). Predict whom one will follow: Followee recommendation in microblogs. In *Proceedings of the international conference on social informatics (Socialinformatics)* (pp. 260–264). IEEE.

Wu, J., Chen, L., Yu, Q., Han, P., & Wu, Z. (2015). Trust-aware media recommendation in heterogeneous social networks. *World Wide Web, 18*(1), 139–157.

Wu, S., Gong, L., Rand, W., & Raschid, L. (2012b). Making recommendations in a microblog to improve the impact of a focal user. In *Proceedings of the sixth ACM conference on recommender systems* (pp. 265–268). ACM.

Xiao, Y., Du, T., Zhu, W., & Li, Q. (2012). Building a tag map for recommendations in microblogging. In *Proceedings of the international conference on management of e-commerce and e-government (ICMECG)* (pp. 169–172). IEEE.

Yamamoto, Y., Kumamoto, T., & Nadamoto, A. (2015). Followee recommendation based on topic extraction and sentiment analysis from tweets. In *Proceedings of the seventeenth international conference on information integration and web-based applications & services* (p. 27). ACM.

Yanardag Delul, P. (2013). Understanding and analysing microblogs. In *Proceedings of the twenty-second international conference on world wide web* (pp. 401–406). ACM.

Yang, X., Guo, Y., Liu, Y., & Steck, H. (2014). A survey of collaborative filtering based social recommender systems. *Computer Communications, 41*, 1–10.

Yazdanfar, N., & Thomo, A. (2013). Link recommender: Collaborative-filtering for recommending URLs to twitter users. *Procedia Computer Science, 19*, 412–419.

Yuan, Z.-m., Huang, C., Sun, X.-y., Li, X.-x., & Xu, D.-r. (2015). A microblog recommendation algorithm based on social tagging and a temporal interest evolution model. *Frontiers of Information Technology & Electronic Engineering, 16*(7), 532–540.

Zhang, J., Fang, Z., Chen, W., & Tang, J. (2015). Diffusion of "following" links in microblogging networks. *IEEE Transactions on Knowledge and Data Engineering, 27*(8), 2093–2106.

Zhao, X., & Tajima, K. (2014). Online retweet recommendation with item count limits. In *Proceedings of the IEEE/WIC/ACM international joint conferences on web intelligence (WI) and intelligent agent technologies (IAT): 1* (pp. 282–289). IEEE Computer Society.