



User interest prediction over future unobserved topics on social networks

Fattane Zarrinkalam^{1,2} · Mohsen Kahani¹ · Ebrahim Bagheri²

Received: 13 November 2017 / Accepted: 28 June 2018 / Published online: 10 July 2018
© Springer Nature B.V. 2018

Abstract

The accurate prediction of users' future interests on social networks allows one to perform future planning by studying how users will react if certain topics emerge in the future. It can improve areas such as targeted advertising and the efficient delivery of services. Despite the importance of predicting user future interests on social networks, existing works mainly focus on identifying user current interests and little work has been done on the prediction of user potential interests in the future. There have been work that attempt to identify a user future interests, however they cannot predict user interests with regard to new topics since these topics have never received any feedback from users in the past. In this paper, we propose a framework that works on the basis of temporal evolution of user interests and utilizes semantic information from knowledge bases such as Wikipedia to predict user future interests and overcome the cold item problem. Through extensive experiments on a real-world Twitter dataset, we demonstrate the effectiveness of our approach in predicting future interests of users compared to state-of-the-art baselines. Moreover, we further show that the impact of our work is especially meaningful when considered in case of cold items.

Keywords User modeling · User interest prediction · Wikipedia category hierarchy · Twitter

1 Introduction

Social networks have shown to be an effective medium for communication and social interaction. Users can interact with others who share similar interests about current trends, topics and world events to communicate news, opinions or other information of interest. As a result of such information sharing and communication processes, in recent years, many researchers have utilized different kinds of information in social networks such as the social relationships, user generated content and temporal aspects in order to identify user interests (Zarrinkalam et al. 2016, 2015; Fani et al. 2016).

✉ Mohsen Kahani
kahani@um.ac.ir

¹ Department of Computer Engineering, Ferdowsi University of Mashhad, Mashhad, Iran

² Laboratory for Systems, Software and Semantics (LS3), Ryerson University, Toronto, Canada

The accurate and timely modeling of user interests can lead to better recommendations, targeted advertising and the efficient delivery of services (Abel et al. 2011b; Asur and Huberman 2010; Kapanipathi et al. 2011). Most existing approaches on social networks focus on modeling users' current interests and little work has been done on the prediction of users' potential future interests. Bao et al. (2013) have proposed a temporal and social probabilistic matrix factorization model to predict future interests of users in microblogging services. They assume that the topic set of the future is known *a priori* and composed only of the set of topics that have been observed in the past, which seems to be an unrealistic limiting assumption, because user's topics of interest on social networks can rapidly change in reaction to real world events (Abel et al. 2011a). Therefore such an approach cannot predict user interests with regard to new topics since these topics have never received any feedback from users in the past (cold item problem) (Bobadilla et al. 2013).

In contrast and in this paper, we aim to extend the state of the art to be able to predict users' interests with regard to future unobserved topics, where both the user interests and the topics themselves are allowed to vary over time. This allows one to perform future planning by studying how users will react if certain topics emerge in the future.

For instance, we are interested in determining whether a given user would be interested in following the news about upcoming *Lady Gaga's Concert* in Canada or not. As such, we are primarily interested in predicting how users would be interested in future topics on Twitter that have not been observed in the past. To address this problem, we propose a prediction framework that works on the basis of temporal evolution of each individual user's interests and utilizes semantic information from knowledge bases such as Wikipedia to predict user future interests and overcome the cold item challenge. Our work is based on the intuition that although it is possible that the topics of interest to the users dramatically change over time as influenced by real-world trends (Abel et al. 2011a), users tend to incline towards topics and trends that are semantically or conceptually similar to a set of core interests. We model high-level interests of users in order to be able to predict user interests over an unobserved set of topics in the future. The key contributions of this paper can be summarized as follows:

- We propose a holistic approach to model high-level interests of users over Wikipedia category hierarchy by generalizing their short-term interests that have been observed over several time intervals.
- We propose a framework that works on the basis of temporal evolution of user interests and considers the semantics of the topics derived from Wikipedia to predict potential future interests of users on Twitter.
- We investigate the influence of different factors such as considering hierarchical structure of Wikipedia category hierarchy and temporal features on the quality of future interest prediction. We compare our model with the state of the art methods which tackle cold item problem and show that our approach achieves improved performance.
- As another contribution, we investigate whether it is possible to predict user interests over those topics which are not observed in the past. This is especially important when considering the fact that user interests change rapidly as they are influenced by real-world events. In our experiments, we will additionally show that the impact of our approach is especially meaningful when considered in such contexts (i.e. cold item problem).

The rest of this paper is organized as follows: In Sect. 2 we describe the related work. Sections 3 and 4 are dedicated to the problem definition and the presentation of the details

of our proposed approach. Section 5 presents the details of our experimental work. Finally, Sect. 6 concludes the paper.

2 Related work

In this paper, we propose a model to predict future interests of users in social networks. It works on the basis of temporal modeling of user historical activities and utilizes Wikipedia to predict future interests of users. Therefore, our work is related to *user interest modeling and prediction* in social networks, *temporal user modeling and recommendation*, and *recommendation based on knowledge bases*. In this section, we review the related works in these three areas.

2.1 User interest modeling and prediction

There is a rich line of research on user interest detection from social networks through the analysis of user generated textual content. To represent user interests, some early works either use *Bag of Words* or *Topic Modeling* approach (Abdel-Hafez and Xu 2013). In the *Bag of Words* approach, user interests are represented as a set of terms extracted from the user contents (Chen et al. 2010; Yang et al. 2012). *Topic Modeling* approach provides a probabilistic model based on the term frequency and term co-occurrences in documents of a given corpus. This approach forms topics by extracting groups of co-occurring terms and views each document as a mixture of various topics (Blei 2012). Latent Dirichlet Allocation (LDA), as a well-known topic modeling method, has been frequently used for interest detection (Ramage et al. 2010; Weng et al. 2010; Xu et al. 2011).

Since the Bag of Words and Topic Modeling approaches focus on terms without considering their semantics and the relationship between them, they do not necessarily utilize the underlying semantics of textual content (Varga et al. 2014; Michelson and Macskassy 2010; Kapanipathi et al. 2014). Furthermore, these approaches assume that a single document contains rich information, as a result they may not perform so well on short, noisy and informal texts like Twitter posts (Cheng et al. 2014; Sriram et al. 2010; Mehrotra et al. 2013). To address these issues, *Bag of Concepts* approach utilizes external knowledge bases to enrich the representation of short textual content and model user interests through semantic entities (concepts) linked to external knowledge bases such as DBpedia and Freebase. Since these knowledge bases represent entities and their relationships, they provide a way of inferring the underlying semantics of content (Michelson and Macskassy 2010; Kapanipathi et al. 2014; Lu et al. 2012).

For example, Abel et al. (2011b, c) have proposed to enrich Twitter posts by linking them to related news articles and then extracting the concepts mentioned in the enriched posts using Web services provided by OpenCalais. The identified concepts are then used to build user interest profiles. Similarly, Kapanipathi et al. (2011) have modeled users' interests by annotating their tweets with DBpedia concepts, and have used these annotations for the purpose of filtering tweets. Song et al. (2015) have proposed a multi-source multi-task schema for user interest inference from social networks in which each user interest is represented by a single concept which is listed in users' LinkedIn profiles.

Most of the works that use the bag of concepts approach for representing user interests fall short when user interests do not necessarily have an exact corresponding concept in the knowledge base. In other words, these models are successful to the extent that they find a

suitable concept or concepts to represent a user interest but they will not be able to model and detect users' interests if such interests are not formally represented in the knowledge base. For instance, when the Wikileaks topic first picked up on Twitter, it was received by a tremendous number of users posting about it; however, at the time, no corresponding DBpedia concept or Wikipedia article was available to link this topic to. In this paper, we follow the earlier works (Zarrinkalam et al. 2015, 2016; Abel et al. 2011a), for extracting user interests where each topic of interest is viewed as a conjunction of several semantic entities which are temporally correlated on Twitter. Therefore, even if a corresponding semantic entity is not available in the external knowledge base, we automatically construct it using existing semantic entities.

As an extension to bag of concepts approach, some researchers extract high-level interests of users by utilizing Wikipedia category graph. For example, Michelson and Macskassy (2010) have proposed a model which first extracts a set of Wikipedia entities from a user's tweets and then identifies the high-level interests of the user by traversing and analyzing the Wikipedia categories of the extracted entities. Kapanipathi et al. (2014) have first extracted weighted primitive interests of a user as a bag of concepts by extracting the entities mentioned in the user's tweets. Then, the high-level interests of the user are extracted by propagating the user interests over the Wikipedia category hierarchy using a spreading activation algorithm where active nodes are initially the set of primitive interests. Similarly, Faralli et al. (2017) and Piao and Breslin (2017) have recently utilized Wikipedia category graph to create a category-based user interest profile. However, because they have focused on passive users, they have extracted primitive interests of a user by linking the user's followee information to the corresponding Wikipedia entities, instead of using the entities mentioned in her tweets. Inspired by these works, we utilize Wikipedia category hierarchy as a knowledge base in order to represent high-level interests of users.

While existing works mainly focus on extracting the current interests of users from social networks, little work has been done on predicting users' future interests. Bao et al. (2013) have proposed a temporal and social probabilistic matrix factorization model that utilize users' sequential interest matrices (user-item matrices) at different time points and the user friendship matrices to predict user interests in the near future in microblogging services. Their work is very similar to ours in a sense that we both try to predict future user interests in microblogging services by taking into account the temporal evolution of user interests. However they are limited by the fact that they assume the set of topics of the future is known *a priori* and composed only of the set of topics that have been observed in the past. Therefore they cannot predict user interests with regard to new topics since these topics have never received any feedback from users in the past (cold item problem). Further, based on the fact that trending topics on microblogging services can rapidly change in reaction to real world events (Abel et al. 2011a), users may be interested in diverse topics over time. As a result, applying the solution of collaborative filtering on the derived sparse data is challenging.

Despite the work proposed in Bao et al. (2013), we have proposed a content-based recommender system that utilizes Wikipedia to predict user interests over an unobserved set of topics in the future, where both user interests and the topics themselves are allowed to vary over time. We have presented our preliminary study of future interest prediction and illustrated its functionality as a proof of concept in Zarrinkalam et al. (2017). This paper extensively extends our previous work with the following improvements: (1) we provide a more comprehensive analysis and review of related work; (2) to model high-level interests of users over Wikipedia category hierarchy, we propose a holistic framework in which different spreading functions and mapping functions are explored; and (3) more

comprehensive experiments are conducted and new findings are reported. Specifically, we studied the influence of different factors on the quality of future interest prediction and investigated the effectiveness of our proposed method to overcome the cold item problem in future interest prediction compared to the state of the art.

2.2 Temporal user modeling and recommendation

Based on the fact that users' interests change over time, temporal aspects have been widely used in conventional recommendation and user modeling systems (Yin et al. 2015; Akbari et al. 2017). One basic solution is to capture users' online behavior over time and build user profiles at different time intervals. For example, Koren (2010) has proposed timeSVD++ to predict movie ratings for Netflix by modeling user temporal dynamics. Bao et al. (2013) have incorporated the time factor into a social matrix factorization model by representing user's dynamic interests as a series of interest matrices (user-item matrices) at different time intervals. Yin et al. (2015) have captured users' changing interests by proposing a unified probabilistic model that extracts both user-oriented topics as intrinsic interests of users and time-oriented topics as temporal context that attracts public attention. Similarly, Gao et al. (2017) have built user interest profiles in different time intervals and have proposed a time-aware item recommender system which captures the evolution of both user interests and item content information via topic dynamics. Liang et al. (2017) have proposed a dynamic user clustering topic model in which users' interests are modeled in each time interval based on both their newly generated posts in that time interval and their extracted interests on previous time intervals.

There is another line of work that builds both long-term and short-term interests of users to handle temporal changes in user interests. For example, Xiang et al. (2010) have proposed a Session-based Temporal Graph (STG) to simultaneously model user short-term and long-term interests over time. Based on the STG model, for the purpose of temporal recommendation, the authors have proposed an algorithm named Injected Preference Fusion (IPF) and extended the personalized random walk on STG. Song et al. (2016) have proposed a multi-rate temporal deep neural network based architecture that captures user interests at different granularity (daily and weekly) to model long-term and short-term interests of users simultaneously. Sang et al. (2015) have proposed a unified probabilistic framework for temporal user modeling on Twitter that simultaneously models user short-term and long-term interests.

In the context of news recommendations, Li et al. (2014) have modeled long-term and short-term interests of users. For modeling user interests, they have first segmented the user's history to several time intervals and build user interest profiles in each time interval. Then, they have used a time-sensitive weighting scheme which is a monotonic decreasing function to extract user's general interests as her long-term interests and the user's current interests as her short-term interests. Similarly, to improve the quality of search personalisation, Vu et al. (2015) have proposed a temporal modeling of user interests based on their clicked documents on the web. To do so, they have used LDA to extract topics from clicked documents of all users. Then to model interests of a user over these topics, given her relevant document set, they have used an exponential decay function as a temporal weighting schema based on which the more recent relevant documents have more weights compared to the distant ones. Based on this approach, by considering different historical user activity time windows, they have modeled three user profiles for each user namely long-term profile, daily profile and session profile.

Given earlier works in the field of temporal user modeling and recommendation show that user interests change over time and tracking user short-term and long-term interests is important for accurate user modeling and performing effective recommendation, our proposed framework works on the basis of temporal modeling of users' historical activities. It is important to note that our main goal in this paper is to propose a new approach to predict users' interests over an unobserved set of topics in the future (cold item problem), not a new temporal user interest modeling approach. Therefore, similar to Sang et al. (2015), to model temporal characteristics of user behavior in microblogging services, we divide the historical time period into different time intervals and model short-term interests of users in each time interval over the discovered active topics in that interval. Therefore, to evaluate our prediction model, we use the same temporal user interest modeling for all the comparison methods which means only the prediction models are compared and evaluated.

It should be noted that in order to be able to predict future interests of users, we incorporate long-term user interests in our prediction model. Given a set of short-term interests of users over active topics in different time intervals, unlike the approach proposed in Sang et al. (2015), we model long-term interests of users over the Wikipedia category hierarchy, which allows us to represent users' long-term interests at different levels of granularity and predict user interests over potential future topics, which may not be observed in the past.

2.3 Recommendation based on knowledge bases

Collaborative Filtering (CF) is one of the most successful and widely used methods to build recommender systems (Ekstrand et al. 2011). However, this approach falls short in recommending new items, which have never received feedback from the users in the past (cold item cases). Further, they usually do not achieve good performance on sparse data (Bobadilla et al. 2013). To overcome these limitations, given the fact that knowledge bases such as YAGO, DBpedia, Microsoft's Satori and Google's Knowledge Graph can provide rich semantic information including both structured and unstructured data, recently, the use of such knowledge bases for augmenting traditional algorithms in the context of content-based or hybrid recommender systems has been the subject of increasing attention (Nakatsuji et al. 2012; Nguyen et al. 2015).

For example, Passant (2010) has proposed dbrec, a music recommender system that computes the semantic distance between music artists based on DBpedia content and recommends the closest artists to a user based on her prior listening habits. Similarly, Di Noia et al. (2012) have proposed a content-based recommendation system in which a new movie is recommended to a user through measuring the similarity between the new movie and the movies rated by the user in the past by utilizing the DBpedia properties. Ostuni et al. (2013) have proposed a hybrid recommender system, named Sprank that expands the work proposed by Di Noia et al. (2012) by mining path-based features that link the past ratings of users together. Similarly, Zhang et al. (2016) have presented a hybrid recommender system that integrates collaborative filtering with different semantic representations of items derived from a knowledge base. The authors have extracted semantic representation of items from structural content, textual content and visual content of Microsoft's Satori knowledge base.

Most of the above work use a knowledge base, such as DBpedia, to calculate the similarity of items; however, they do not integrate hierarchical structure of the knowledge base in their recommendations. Cheekula et al. (2015) have proposed a content-based recommendation method that utilizes user interests modeled over Wikipedia category hierarchy to

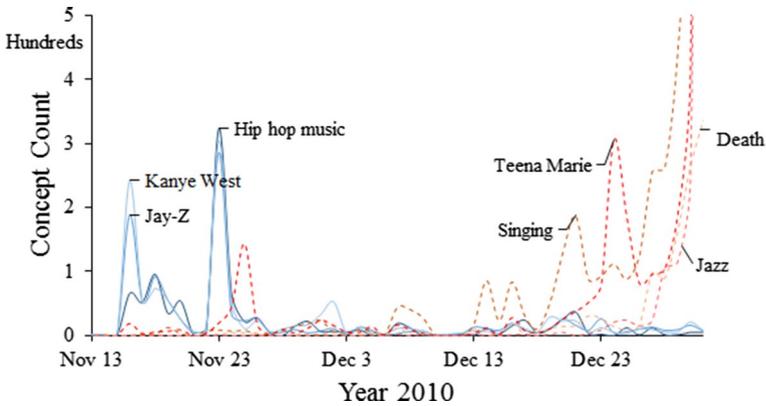


Fig. 1 Occurrence frequencies of related entities to two real world topics/events, *Jay-Z and Kanye West's performance* and *Teena Marie's death*

recommend relevant entities. To build hierarchical user interests, they follow the approach in Kapanipathi et al. (2014). Our work is similar to them in a sense that both works model high-level interests of users over the Wikipedia category hierarchy. However, they overlook the evolution of user interests over time in their user model. Further, our main goal is different. We aim at predicting user interests over unobserved topics in the future.

3 Problem definition

The objective of our work is to predict user interests with regard to unobserved topics of the future on Twitter. To be able to achieve this goal, we rely on temporal and historical user interest profile information in order to predict how users would react to potential topics of the future.

Recent studies have shown that trending topics on social networks can rapidly change in reaction to real world events and therefore, the set of topics can change between different time intervals (Abel et al. 2011a; Huang et al. 2017). For example, in Fig. 1, we show how the occurrence frequency of important entities, which are related to two topics/events named *Jay-Z and Kanye West's performance* and *Teena Marie's death* change over time. The occurrence frequencies of these entities show that *Jay-Z and Kanye West's performance* topic reaches its peaks in mid-November and then decreases rather slowly over the next weeks, while the latter becomes trendy in late December. It is noted that topic and event are used interchangeably in this paper.

In the context of future interest prediction on microblogging services, there have already been work in the literature that focus on how users' interests would evolve over time (Bao et al. 2013); however, such work assume that the set of topics stay the same over time. This seems to be an unrealistic limiting assumption as it fails to capture the natural evolution of topics in time and consequently not able to predict future interests of users to a set of topics that have not been observed in the past.

Our work extends the state of the art by modeling topical interests of a user where both (1) the users' association with topics and (2) the topics themselves are allowed to change

over time, thus allowing the prediction of user interests based on historical data over an unobserved set of topics in the future.

To this end and to express the temporal dynamics of topics and user interests, we divide the users' historical activities into L discrete time intervals $1 \leq t \leq L$ and then for each time interval t , we extract a set of topics and the distribution of user interests over those topics, using the microposts which are published in the corresponding time interval. More formally, let t be a specified time interval and $Z^t = \{z_1, z_2, \dots, z_K\}$ be K active topics in time interval t , which is not necessarily the same as the topics in the previous or next time intervals, for each user $u \in \mathbb{U}$, we define her topic profile in time interval t over Z^t , denoted as $TP^t(u)$, as follows:

Definition 1 (*Topic Profile*) The topic profile of user $u \in \mathbb{U}$ in time interval t , with respect to Z^t , denoted by $TP^t(u)$, is represented by a vector of weights over the K topics, i.e. $(f_u^t(z_1), \dots, f_u^t(z_K))$, where $f_u^t(z_k)$ denotes the degree of u 's interest in topic $z_k \in Z^t$. A user topic profile is normalized using L1-norm.

It is worth noting that, because a user topic profile is the specification of the user's interests towards active topics in a given time interval, it actually represents the short-term interests of the user which leads to capturing user preferences both in a timely fashion and at a fine-grained level (Sang et al. 2015). In the paper, we use user topic profile and user's short-term interests, interchangeably.

After modeling historical user activities in each time interval $t : 1 \leq t \leq L$, as a result of which each user will have L user topic profiles, one for each of the time intervals, our objective is to predict future interests of users as follows:

Definition 2 (*Future Topic Profile*) Given the topic profiles for a user u in each time interval of the historical time period, i.e. $TP^1(u), \dots, TP^L(u)$, as well as a set of topics in time interval $L + 1$, Z^{L+1} which might not have been observed in the previous time intervals, we aim to predict $\hat{TP}^{L+1}(u)$ that we refer to as the future topic profile of user u towards Z^{L+1} .

Figure 2 shows the overview of our approach to address the challenge defined in Definition 2. As depicted in Fig. 2, we divide this problem into two sub-problems: *Temporal Modeling of User Historical Activities* and *Future Interest Prediction*, in which the output of the first sub-problem becomes the input of the second one. In the following section, we describe our proposed approach for addressing these two sub-problems in detail.

4 Proposed approach

In this section, we first introduce our approach for temporal modeling of user historical activities and then based on the output of this step, we describe our prediction method to predict future interests of users.

4.1 Temporal modeling of user historical activities

As explained earlier, to express the temporal dynamics of topics and user interests, we divide the users' historical activities into L discrete time intervals $1 \leq t \leq L$ and extract L topic sets Z^1, Z^2, \dots, Z^L and the topic profile of each user u within these time intervals

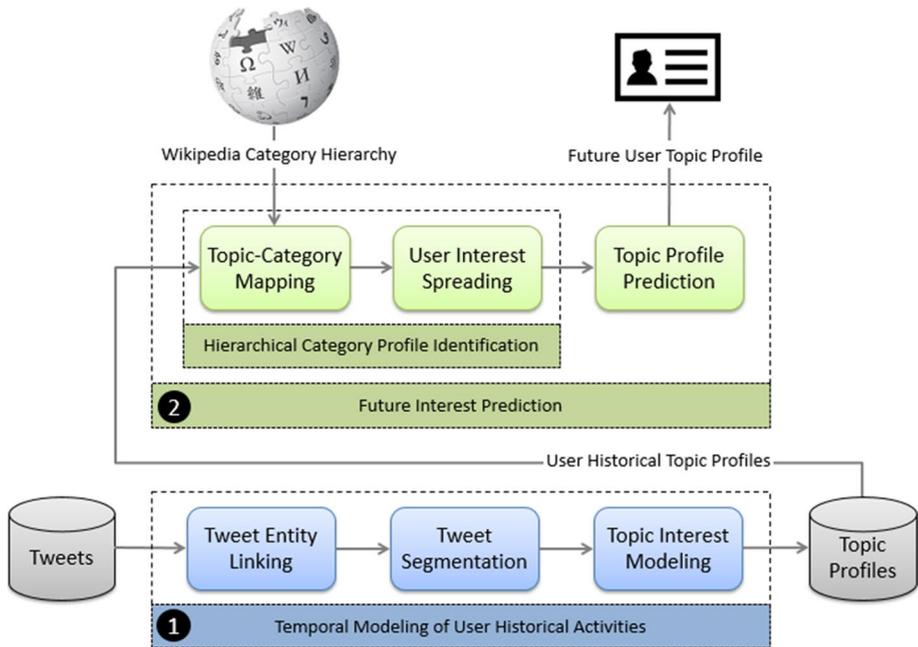


Fig. 2 Overview of our proposed approach

as $TP^1(u), \dots, TP^L(u)$. It should be noted that topic/event and user interest detection methods from microblogging services have already been studied in the literature and therefore are not the focus of our work and we are able to work with any topic/event and user interest detection methods to extract Z^t and $TP^t(u)$.

Considering M^t , the set of microposts published in time interval t , it is possible to extract topics Z^t using topic modeling methods. LDA (Blei et al. 2003) is one of the well-known unsupervised techniques used for identifying latent topics from a corpus of documents. However, because it is designed for regular documents not microposts, it may not perform well on short, noisy and informal texts like tweets and might suffer from the sparsity problem (Cheng et al. 2014; Mehrotra et al. 2013; Sriram et al. 2010). To address this issue, we employ Twitter Latent Dirichlet Allocation (Twitter-LDA) (Zhao et al. 2011), which is a topic modeling approach specifically for Twitter content to extract topics. Twitter-LDA, as an extension to the standard LDA, assumes that each tweet is produced by a single topic and a background model.

Furthermore, as proposed in Zarrinkalam et al. (2015, 2016), we enrich each micropost m from our corpus M^t by using an existing semantic annotator and employ the extracted entities, instead of words, within the topic detection process. Therefore, in our work, each micropost is considered to be a set of one or more semantic entities that collectively denote the underlying semantics of the microposts. This can not only lead to the reduction of noisy content, but also provides a semantic representation of topics which is a requirement for our proposed model. As explained later in the experiments section, we employ TAGME (Ferragina and Scaiella 2012) to link microposts to Wikipedia entities. For instance, for a tweet such as "Pete Rock Says Kanye, Jay-Z

Collabo 'The Joy' Was Made In '97", by using TAGME (Ferragina and Scaiella 2012), we model it as a collection of three semantic entities, namely *Pete_Rock*, *Kanye_West* and *Jay-Z*.

Therefore, we view a topic, defined in Definition 3, as a distribution over Wikipedia entities. In other words, let $\mathbb{E} = \{e_1, e_2, \dots, e_{|\mathbb{E}|}\}$ be a set of Wikipedia entities extracted from \mathbb{M}^t , we consider \mathbb{E} to be the vocabulary in the topic detection method. A topic can, more formally, be defined as follows:

Definition 3 (*Topic*) Let \mathbb{M}^t be a corpus of microposts published in time interval t and \mathbb{E} be the vocabulary of Wikipedia entities, a topic in time interval t , is defined to be a vector of weights, i.e. $z = (h_z(e_1), \dots, h_z(e_{|\mathbb{E}|}))$, where $h_z(e_i)$ shows the participation score of entity $e_i \in \mathbb{E}$ in forming topic z . Collectively, $\mathbb{Z}^t = \{z_1, z_2, \dots, z_K\}$ denotes a set of K topics extracted from \mathbb{M}^t .

Given Wikipedia entities \mathbb{E} forming the vocabulary space, by applying Twitter-LDA over the microposts \mathbb{M}^t , it produces the following two artifacts:

1. K topic-entity distributions, where each topic entity distribution associated with a topic $z \in \mathbb{Z}^t$ represents active topics in \mathbb{M}^t , i.e. $z = (h_z(e_1), \dots, h_z(e_{|\mathbb{E}|}))$
2. $|\mathbb{U}|$ user-topic distributions, where each user-topic distribution associated with a user u represents the topic profile of user u in time interval t , i.e. $TP^t(u) = (f_u^t(z_1), \dots, f_u^t(z_K))$.

Now, with the above formulation, given a corpus of microposts \mathbb{M} , we will break it down into L intervals and perform the above process separately on each of the intervals. This will produce L topic profiles for every user u in our user set, i.e. $TP^1(u), \dots, TP^L(u)$, which is the required input for our future user interest profile prediction problem defined in Definition 2.

4.2 Future interest prediction

Given $TP^1(u), \dots, TP^L(u)$, our goal is to predict potential interests of each user $u \in \mathbb{U}$ over \mathbb{Z}^{L+1} . It is important to point out that since $L + 1$ is in the future, the topics \mathbb{Z}^{L+1} may have not been observed.

Our prediction model is based on the observation that, although short-term user interests are driven by the shifts and changes in real world events and trends, they tend to stay consistent in long-term intervals (Sang et al. 2015). Simply put, while user's short-term interests might change over time, they revolve around some fundamental concepts for each user.

For example, assuming users' historical activities can be divided into three discrete time intervals (i.e. t_1, t_2, t_3) where t_4 would be the future time interval, Table 1 gives information about four topics that a given user, who we call @mary, is interested in each of the time intervals. Topic $z_1 \in \mathbb{Z}^1$ refers to the selling of Beatles songs on iTunes, topic $z_2 \in \mathbb{Z}^2$ refers to the hip-hop music collaboratively produced by American rappers Jay-Z and Kanye West, topic $z_3 \in \mathbb{Z}^3$ represents Lady Gaga's concert in Canada and topic $z_4 \in \mathbb{Z}^4$ is related to Teena Marie's death. It can be observed that although @mary's short-term interests has changed over time, all of these three topics are related to the generic category of `Category:Music`. Thus her short-term interests stay consistent over time.

As a result, in order to be able to achieve predictability over future topics, we generalize individual user's short-term interests that have been observed over several time intervals to

Table 1 Four topics that @mary is interested in within different time intervals and their related Wikipedia categories

Topic#	Top-5 entities	Wikipedia categories	Time
z_1	The_Beatles, Music, iTune, Album, rock	English_rock_music_groups, Songs, Apple_Records_artists, British_psychedellic_rock_music_groups, Apple_Inc_services	t_1
z_2	Mixtape, Hip_hop_music, Rapping, Kanye_West, Jay-Z, Remix	20th-century_American_musicians, American_hip_hop_record_producers, African-American_music, Brit_Award_winners, Grammy_Award_winners	t_2
z_3	Lady_Gaga, Song, Album, Concert, Canadian_Hot_100	Music_performance, American_female_pop_singers, American_n_female_musicians, Songs, Grammy_Award_winners	t_3
z_4	Teena_Marie, Death, Jazz, singing, Florida	2010_deaths, African-American_music, Death, American_dance_musicians, American_female_singers	t_4

gain a good insight into the user's high-level interests, i.e. long-term interests. It not only will allow us to generalize users' interests, but also enables us to transfer users' interests across different time intervals that do not necessarily have the same set of topics. To this end, given the topic profile of each user over different time intervals, we first utilize Wikipedia category structure for modeling long-term interests of users (*Hierarchical Category Profile Identification*). Then, the inferred category profile of the user from the previous step is used to predict future interests of the user (*Topic Profile Prediction*).

4.2.1 Hierarchical category profile identification

In this section, we aim at utilizing the Wikipedia category hierarchy to generalize the topic-based representation of user interests into a category-based representation for modeling long-term interests of users. There are two main reasons behind this choice: (1) it has been experimentally shown in Siehndel and Kawase (2012) that categories of interest to every user stay approximately stable over time. As a result, they can be considered to be a suitable long-term user interest representation. (2) exploiting hierarchical structure of categories provides a flexible approach through which user interests are able to be represented at different levels of granularity (Kapanipathi et al. 2014).

A major challenge in utilizing Wikipedia category structure as a hierarchy is that, it is a cyclic graph instead of a strict hierarchy. This is because categories in Wikipedia are created and edited collaboratively by many different users where any user is free to create or link categories to each other; hence, potentially leading to cyclic references between categories. Using Wikipedia categories without removing cycles would be problematic because cycles make it non-trivial to determine the hierarchical relationships between categories (Boldi and Monti 2016; Kapanipathi et al. 2014). Therefore, as a pre-process in our user modeling approach, we first transform the Wikipedia category structure into a hierarchy by adopting the approach proposed in Kapanipathi et al. (2014). To do so, we first remove the following Wikipedia admin categories: {wikipedia, wikiprojects, lists, mediawiki, template, user, portal, categories, articles, pages} (Ponzetto and Strube 2007). Then, to transform the Wikipedia category structure into a hierarchy, we select `Category:Main_Topic_Classifications`, which subsumes 98% of the categories as the root node of the hierarchy. Next, we assign the abstract level of each category based on its shortest path length to the root node. As the last step, we remove all the directed edges from a category of lower hierarchical level (specific) to a higher hierarchical level (conceptually abstract). The output of this process is a Wikipedia category hierarchy (*WCH*), a directed acyclic graph whose nodes are the Wikipedia categories \mathbb{C} with an edge from $c_i \in \mathbb{C}$ to $c_j \in \mathbb{C}$ whenever c_j is a subcategory of c_i .

Next, given the topic profiles of user u in L time intervals, and Wikipedia category hierarchy, we build hierarchical category profile of user $u \in \mathbb{U}$ over the L historical time intervals, as defined in Definition 4, based on Algorithm 1.

Algorithm 1: Modeling hierarchical category profile for user u

```

1 Input:  $TP^1(u), \dots, TP^L(u)$  and Wikipedia category hierarchy ( $WCH$ )
2 Output: Hierarchical Category Profile of user  $u$ ,  $HCP^L(u)$ 
3 Initialization:  $WCH^0(u)$  is a copy of  $WCH$  in which the value of each node is set to zero
   1: for  $TP^t(u) = (f_u^t(z_1), \dots, f_u^t(z_K)) : 1 \leq t \leq L$  do
   2:    $WCH^t(u) \leftarrow WCH^{t-1}(u)$ 
   3:   for  $z \in \mathbb{Z}^t$  do
   4:     if  $f_u^t(z) > 0$  then
   5:        $\mathbb{C}_z = \{(c, \Phi(c, z)) | c \in \mathbb{C}\} \leftarrow \text{MapTopicsToCategories}(z)$ 
   6:        $\mathbb{C}_z \leftarrow \text{RemoveRedundancy}(\mathbb{C}_z)$ 
   7:       for  $(c, \Phi(c, z)) \in \mathbb{C}_z$  do
   8:          $av_u^t(c) \leftarrow f_u^t(z) \times \Phi(c, z)$ 
   9:          $WCH^t(u) \leftarrow \text{SpreadUserInterests}(c, av_u^t(c), WCH^t(u))$ 
10:  $HCP^L(u) \leftarrow WCH^L(u)$ 
11: return  $HCP^L(u)$ 

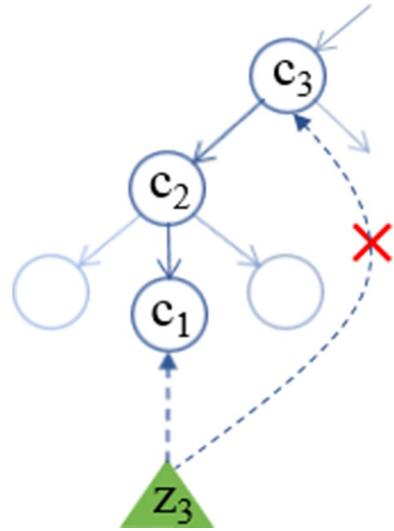
```

Definition 4 (*Hierarchical Category Profile*) Given the Wikipedia category hierarchy, WCH , the hierarchical category profile of user $u \in \mathbb{U}$ over the L historical time intervals, called $HCP^L(u)$, is represented by the same structure as WCH , in which the weight of each node c , denoted as $g_u^L(c)$, denotes the degree of u 's interest in category $c \in \mathbb{C}$ over the L historical time intervals.

As shown in Algorithm 1, to build the hierarchical category profile of a user u , we first make a copy of WCH for user u and initialize the weight of all nodes to zero, called $WCH^0(u)$. Then, for each time interval t , given the topic profile of user u in that time interval, $TP^t(u) = (f_u^t(z_1), \dots, f_u^t(z_K))$, for every topic $z \in \mathbb{Z}^t$ that user u is interested in, i.e. $f_u^t(z) > 0$, we infer the related categories of z , via a mapping function described in the *Topic-Category Mapping* section. The mapping function returns a set of pairs $(c, \Phi(c, z))$ such that c indicates a Wikipedia category related to topic $z \in \mathbb{Z}^t$ and $\Phi(c, z)$ denotes their degree of relatedness (See Line 5). Therefore, in this step, each topic z of interest to user u in time interval t is mapped to a set of Wikipedia categories, called \mathbb{C}_z . Table 1 illustrates 5 related Wikipedia categories for each of @mary's topics of interest.

We assume that when a user is interested in a certain category, such as `Category:American_female_pop_singers`, she might also be interested in broader related categories, e.g. `Category:American_female_singers`. Therefore, we propagate user's categories of interest over the Wikipedia hierarchy. However, as the categories in \mathbb{C}_z could either be generic or specific, they may belong to different levels of the Wikipedia category hierarchy and consequently, it is possible for the topic z to activate a generic category redundantly through both direct and hierarchical relationships. To solve this problem, before applying the spreading activation function, we remove a generic category $c \in \mathbb{C}_z$, if there is a specific category in \mathbb{C}_z which can activate c through the propagation process. This is achieved by calling the *RemoveRedundancy*(\mathbb{C}_z) function in Line 6. For example, as mentioned in Table 1, for topic z_3 , i.e. *Lady Gaga's concert*, \mathbb{C}_{z_3} includes both `Category:American_female_pop_singers` and `Category:American_female_musicians`. As it is shown in Fig. 3, in the Wikipedia category hierarchy, $c_3 : \text{Category:American_female_musicians}$ is the parent of $c_1 : \text{Category:American_female_pop_singers}$. Thus, through

Fig. 3 The output of RemoveRedundancy function for the topic of *Lady Gaga's concert* (topic z_3). Here, c_1 denotes Category:American_female_pop_singers, c_2 denotes Category:American_female_singers and c_3 denotes Category:American_female_musicians



$RemoveRedundancy(\mathbb{C}_{z_3})$, we remove c_3 , which will be activated through the propagation process from c_1 .

Given $\mathbb{C}_z = \{(c, \Phi(c, z)) | c \in \mathbb{C}\}$ as our initially activated category nodes in the propagation process, the activation value of each category $c \in \mathbb{C}_z$, which is the degree of interest of user u to category c at time interval t , denoted as $av_u^t(c)$, is calculated by multiplying the degree of interest of user u to topic $z \in \mathbb{Z}^t$, $f_u^t(z)$, and the relatedness value of topic z to category c , $\Phi(c, z)$ (see Line 8).

Next, given an activated category node and its activation value, we utilize the *spreadUserInterests* function to infer more abstract and broader user interests by exploiting the Wikipedia category hierarchy (see Line 9). We repeat the above described steps for the L topic profiles of user u . The final result that is $WCH^L(u)$ represents Hierarchical category profile of user u over the L historical time intervals, denoted as $HCP^L(u)$.

For example, Given the topics of interest to @mary in time intervals t_1 , t_2 and t_3 , Fig. 4 shows the process of spreading her interests over Wikipedia category hierarchy to build the hierarchical category profile of @mary over these three time intervals, i.e. $HCP^3(@mary)$. We now explain in more detail, how the two main functions in Algorithm 1, i.e. the mapping and spreading functions, operate.

Topic-Category Mapping The objective of this section is to map a given topic to a set of semantically related Wikipedia categories. As mentioned in Sect. 4.1, we view each topic as a distribution over Wikipedia entities \mathbb{E} which are already associated with one or more categories in Wikipedia. Therefore, we use two possible approaches to calculate the degree of relatedness of a topic z to a Wikipedia category $c \in \mathbb{C}$, which are based on the constituent entities of the topic: (1) Attribution approach, and (2) Embedding approach.

In the *Attribution* approach, for a given topic z , those categories in Wikipedia that are directly associated with the constituent entities of the topic are considered as related categories. For instance, topic z_3 which refers to *Lady Gaga's concert*, is related to Category:American_female_pop_singers, which is one of the categories associated with the constituent entities of the topic, i.e. *Lady_Gaga*.

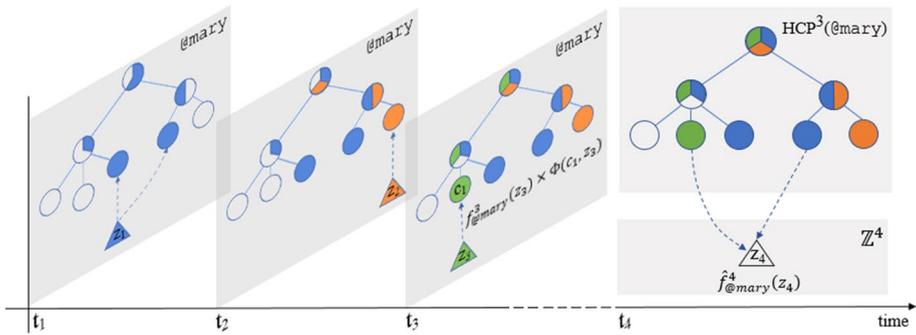


Fig. 4 The process of spreading @mary’s interests over Wikipedia category hierarchy to build $HCP^3(@mary)$ and predicting her interests with regard to topic z_4 which represents a potential topic in future (t_4)

Based on this approach, we can calculate the degree of relatedness of topic $z = (h_z(e_1), \dots, h_z(e_{|\mathbb{E}|}))$ to category $c \in \mathbb{C}$, called $\Phi_a(z, c)$, as follows:

$$\Phi_a(z, c) = \sum_{i=1}^{|\mathbb{E}|} h_z(e_i) \times \delta_c(e_i) \tag{1}$$

Here, $\delta_c(e)$ is set to 1 if entity $e \in \mathbb{E}$ is a Wikipedia entity that has the Wikipedia category c , otherwise, it is zero. Further, $h_z(e)$ is the distribution of entity e in topic z , calculated as described in Sect. 4.1.

The objective of the *Embedding* approach is to utilize a distributed representation for Wikipedia entities and categories that can capture their semantic relatedness. There has been a growing interest in distributed representation of words which states that words that occur in similar contexts are semantically similar. The skip-gram model (Mikolov et al. 2013) is one of the most popular methods to learn word representations so that the semantic similarity of words can be measured in terms of geometric distance between the corresponding vectors. Recently, Li et al. (2016) have proposed a framework that simultaneously learns entity and category representations, in the same feature space, from Wikipedia knowledge base. We adopt their Hierarchical Category Embedding (*HCE*) model, which incorporates category hierarchies into category and entity embeddings.

Given the semantic representation of entities from *HCE*, we derive the embedding vector of z , called HCE_z , by averaging all entity vectors corresponding to the constituent entities of z , as follows:

$$HCE_z = \frac{1}{|\mathbb{E}|} \sum_{i=1}^{|\mathbb{E}|} h_z(e_i) \times HCE_{e_i} \tag{2}$$

where HCE_e is a vector representation of entity $e \in \mathbb{E}$ and $h_z(e)$ is the distribution of entity e in topic z .

Finally, the degree of relatedness of topic z to category c , $\Phi_e(z, c)$ is calculated based on the cosine similarity between the embedding vectors of category c , HCE_c , and topic z , HCE_z .

User Interest Spreading Given a category of interest to a user, in this section, we aim at utilizing a spreading activation function to determine her broader interest categories. The spreading activation function takes an activated node and its activation value as input and iteratively spreads the activation value of the corresponding node out to its parent nodes until a stopping condition is reached (e.g., up to the root node). Formally, the basic spreading activation function is defined as follows:

$$av(c_i) = av(c_i) + av(c_j) \times D \quad (3)$$

where c_i is the category to be activated and $av(c_i)$ is its activation value; c_j is a subcategory of c_i ; and D is a decay factor. If a node is activated by multiple nodes, the activation is accumulated in this node.

For the task of user profiling, it has been experimentally shown in Kapanipathi et al. (2014) and Cheekula (2016) that using different constant values as decay factor ($D = 0.4, 0.6, 0.8$), because of propagating the activation values up the hierarchy without considering the distribution and characteristics of the nodes, results in a hierarchy which has higher weights for categories that are broader. This is not appropriate for the purpose of developing reasonably accurate user profiles because it makes generic user interests more important than specific ones and consequently cannot help recommender systems to generate more user-specific recommendations. We alleviate this problem by taking (1) the distribution of categories in each level of Wikipedia category hierarchy and (2) characteristics of each category such as its specificity and priority during the spreading process into account. Therefore, in our experiments, we exploit three functions, namely: (1) Bell Log, (2) Specificity and (3) Priority. These activation functions are introduced in the following.

The *Bell Log* approach (Kapanipathi et al. 2014) is based on the idea that the distribution of categories across the Wikipedia hierarchy follows a bell curve and this uneven distribution impacts the propagated activation value by increasing the weight of categories with more sub-categories. Therefore, during the propagation step, it penalizes the activation value of each category based on the number of sub-categories at its child level. The formulation of this spreading function is as follows:

$$av(c_i) = av(c_i) + av(c_j) \times BL(c_i) \quad (4)$$

$$BL(c_i) = \frac{1}{\log_{10} nodes(\ell_i + 1)} \quad (5)$$

where ℓ_i is the hierarchical level of category i and the function $nodes(\ell)$ returns the number of nodes at hierarchical level ℓ .

The *Specificity* approach is based on the idea that more specific user interests are more useful than generic interests for the purpose of user profiling and result in more accurate recommendations (Cheekula 2016). For example, based on Wikipedia category hierarchy, we can infer that `Category:Hip_hop_production` is more specific than `Category:Music`. As a result, estimating that a given user is interested in `Category:Hip_hop_production` yields fewer results and is hence more accurate in the prediction/recommendation step compared to `Category:Music`, as it may be too generic and add more noise in the recommendation results. Thus, we incorporate the specificity of each category in the hierarchy as a parameter in the spreading activation function by penalizing the activation value with respect to the specificity of a category c . The formulation of this spreading function is as follows:

$$av(c_i) = av(c_i) + av(c_j) \times S(c_i) \quad (6)$$

$$S(c_i) = 1 - \frac{\log(|\text{successor}(c_i)| + 1)}{\log N} \tag{7}$$

where $S(c)$ measures the specificity of a category node as the ratio of the number of nodes subsumed by a category to the total number of nodes in the Hierarchy. The function $\text{successor}(c)$ returns all successors of the category node c in the Wikipedia category hierarchy and N denotes the number of all categories in the hierarchy.

The *Priority* approach is based on the idea that different super-categories of a category have varying levels of importance as related to the category and as a result they should not be given equal priority during the propagation step (Cheekula 2016). In other words, the activation value of a category should not spread equally to all its super categories. For example, `Category:Hip_hop_production`, `Category:Vocal_music`, `Category:Lyrics` are three super categories of `Category:Rapping`. However, intuitively, these categories are not equally related to `Category:Rapping`. As such, we prioritize super categories of a category by adding weights to the edges in the hierarchy that represent the similarity between each two categories. The formulation of this spreading function is as follows:

$$av(c_i) = av(c_i) + av(c_j) \times P(c_i, c_j) \tag{8}$$

$$P(c_i, c_j) = \frac{HCE_{c_i} \cdot HCE_{c_j}}{\|HCE_{c_i}\|_2 \|HCE_{c_j}\|_2} \tag{9}$$

where $P(c_i, c_j)$ measures the semantic similarity between two categories c_i and c_j based on the cosine similarity between their embedding vectors. HCE_c denotes the embedding vector of category c , retrieved from the embeddings of the HCE model on Wikipedia (Li et al. 2016).

At the end of this process, for each user $u \in \mathbb{U}$, given her topic profiles in each time interval of the historical time periods, $TP^1(u), \dots, TP^L(u)$, we have built her hierarchical category profile, $HCP^L(u)$, as defined in Definition 4. $HCP^L(u)$ represents the high-level interests of user u and will be the input of the next section to predict how users would react to potential topics of the future.

4.2.2 Topic profile prediction

In this section, our goal is to predict future topic profile of each user over \mathbb{Z}^{L+1} . Formally, given the hierarchical category profile of user u , $HCP^L(u)$, and a set of unobserved topics for time interval $L + 1$, $\mathbb{Z}^{L+1} = \{z_1, z_2, \dots, z_K\}$, we are interested in predicting a topic profile for user u over \mathbb{Z}^{L+1} , $\hat{TP}^{L+1}(u) = (\hat{f}_u^{L+1}(z_1), \dots, \hat{f}_u^{L+1}(z_K))$. We calculate $\hat{f}_u^{L+1}(z)$ as follows:

$$\hat{f}_u^{L+1}(z) = \sum_{i=1}^{|\mathbb{C}|} \Phi(z, c_i) \times g_u^L(c_i) \tag{10}$$

where $\Phi(z, c)$ is a function that calculates the degree of relatedness of topic z to category $c \in \mathbb{C}$ which is the output of mapping function (i.e. $\text{mapTopicsToCategories}(z)$) and $g_u^L(c)$

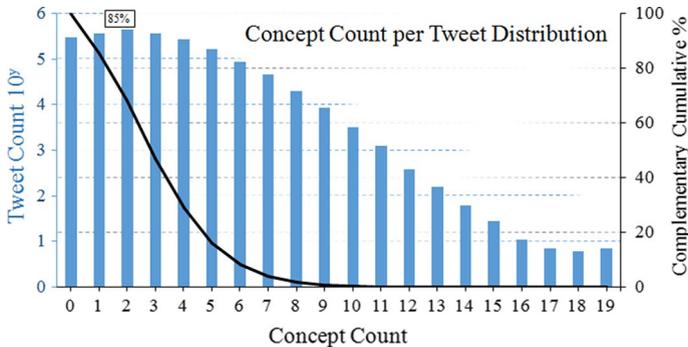


Fig. 5 The number of concepts per tweet and its complementary cumulative distribution

denotes the weight of node c in $HCP^L(u)$ which is the hierarchical category profile of user u over the L historical time intervals.

Given the hierarchical category profile of user @mary over three historical time intervals, i.e. $HCP^3(@mary)$, and topic z_4 as a potential topic in time interval t_4 , Fig. 4 shows the process of future interest prediction to predict the degree of interest of @mary over z_4 , i.e. $\hat{f}_{@mary}^4(z_4)$. For example, by spreading the category of interest of @mary in historical time intervals t_1, t_2 and t_3 , reported in the third column of Table 1, over Wikipedia category hierarchy, it can be inferred that @mary is interested in Category:American_female_singers. Therefore, when Teena Marie who is an American singer dies in time interval t_4 (topic z_4), @mary might potentially be interested to follow the news related to z_4 , which also belongs to the category of Category:American_female_singers.

5 Experiments

In this section, we describe our experiments in terms of the dataset, setup and performance compared to the state of the art.

5.1 Dataset and experimental setup

In our experiments, we use a publicly available Twitter dataset¹ collected and published by Abel et al. (2011b). It consists of approximately 3M tweets posted by 135,731 unique users between November 1 and December 31, 2010. To focus on the active Twitter users for user modeling, in our experiments, we only consider Twitter users who have published more than 100 tweets in our dataset as golden users, which resulted in 2,581 unique users who have collectively posted 2,554,540 tweets.

To annotate the text of each tweet with Wikipedia entities, we adopt the TAGME RESTful API,² which resulted in 350,731 unique entities. The choice of TAGME was motivated by a study that showed this semantic annotator performed reasonably on different types of text such as tweets, queries and Web pages (Cornolti et al. 2013). The distribution of concepts (Wikipedia entities) in a single tweet is shown in Fig. 5. From the tweet content

¹ <http://www.wis.ewi.tudelft.nl/umap2011/>.

² <https://tagme.d4science.org/tagme/>.

perspective, the complementary cumulative distribution of the concepts shows that in more than 85% of the tweets, there is at least one concept.

In the experiments, we divide our dataset into $L + 1$ time intervals. In each time interval $t : 1 \leq t \leq L + 1$, in order to extract active topics, \mathbb{Z}^t , and topic profile of each user u , $TP^t(u)$, we apply the implementation of Twitter-LDA³ on the collection of entities mentioned in the tweets published in t , \mathbb{M}^t , as described in Sect. 4.1. The first L time intervals serve as our training data and the last is employed for testing. It should be noted that to simplify this process, we assume the short-term topic spaces of different time intervals share the same number of topics K . Further, given Twitter-LDA requires the number of topics K to be known a priori, we repeated all of our experiments on different number of topics: 20, 30 and 40.

To prepare the Wikipedia category graph, we downloaded the freely available English version of DBpedia,⁴ which is extracted from Wikipedia dumps dated April 2016. This dataset consists of 1,411,022 categories with 2,830,740 subcategory relations between them. As described in Sect. 4.2.1, we transform the Wikipedia category graph into a hierarchy by adopting the approach proposed in Kapanipathi et al. (2014). The outcome of this process is a hierarchy with a height of 26 and 1,016,584 categories with 1,486,019 links among them. It should be noted that the distribution of categories in different levels of hierarchy follows a bell curve such that most of the topics are covered in levels 7, 8 and 9. This is the idea behind the *Bell log* activation function.

It is important to note that although the Twitter dataset used in our experiments includes tweets published in 2010, the Wikipedia version used for the purpose of tweet entity linking is from 2016. This increases the likelihood of mentions that are in a tweet being covered by corresponding Wikipedia entities and consequently it will improve the quality of temporal user interest modeling when building topic profile of users, compared to utilizing an older version of Wikipedia. However, given all the comparison methods employ the same set of topics and user topic profiles in their prediction model, the impact of the Wikipedia version is the same for all the baselines. Therefore, we believe the version of Wikipedia does not positively favor any of the comparison method and as a result, the impact of considering different versions of Wikipedia for the purpose of tweet entity linking is not investigated in this paper.

5.2 Evaluation methodology and metrics

Given the outputs of Twitter-LDA over $L + 1$ time intervals, as mentioned earlier, we consider the first L extracted topic profiles of each user u , $TP^1(u), \dots, TP^L(u)$, as the historical interests of user u for training and $TP^{L+1}(u)$ as the potential interests of user u over \mathbb{Z}^{L+1} for testing.

We adopt two metrics namely Normalized Discounted Cumulative Gain (nDCG) and Mean Average Precision (MAP) as two well-known metrics in information retrieval for ranking the quality of the results. For a given user u , to calculate these evaluation metrics, it is required to specify the *ground truth* and the *ranked predicted list* of that user. Therefore, in order to build the ground truth of each user u , given $TP^{L+1}(u) = (f_u^{L+1}(z_1), \dots, f_u^{L+1}(z_K))$ that represents the interest profile of user u in time interval $L + 1$ over \mathbb{Z}^{L+1} , we first rank her topics

³ <https://github.com/minghui/Twitter-LDA/>.

⁴ <http://downloads.dbpedia.org/>.

of interest based on the degree of interest of user u to each topic $z \in \mathbb{Z}^{L+1}$, i.e. $f_u^{L+1}(z)$. Then, we consider the top-5 topics of most interest to user u as her *ground truth*. Further, given $\hat{TP}^{L+1}(u) = (\hat{f}_u^{L+1}(z_1), \dots, \hat{f}_u^{L+1}(z_K))$ as the predicted interest profile of user u in time interval $L + 1$, we rank her topics of interest based on the predicted degree of interest of user u toward each topic $z \in \mathbb{Z}^{L+1}$, i.e. $\hat{f}_u^{L+1}(z)$. The resulting ranked list is considered as the *ranked predicted list* for user u .

In our experiments, we also use paired t-test with 95% confidence level for testing the statistical significance of the observed differences between different methods.

5.3 Comparison methods

Our goal is to predict future user interests, i.e. the degree of user interests over certain topics that emerge in the future, which may have not been observed in the past. Among different recommendation strategies, Collaborative Filtering (CF) based methods, which utilize user historical interactions, have gained significant success (Ekstrand et al. 2011). However, CF methods cannot recommend new items since these items have never received any feedback from users in the past, known as the *cold item problem*. As a result, this line of works is not able to address the problem discussed in this paper and consequently we do not consider them as a comparison methods. To tackle the cold item problem, content-based and hybrid approaches that incorporate item content are an appropriate choice (Bobadilla et al. 2013). Thus, we consider the following methods as baselines.

It is important to note that all the comparison methods employ the same set of topics $\mathbb{Z}^1, \dots, \mathbb{Z}^L$ and user topic profiles $TP^1(u), \dots, TP^L(u)$, the output of temporal user modeling described in Sect. 4.1, in their prediction model. In other words, only the prediction model is different in each approach.

SCRS Based on the method proposed by Di Noia et al. (2012), user’s interests in the future are semantically similar to the ones a user has been interested in the past. They use existing knowledge bases to extract item features to compute the similarity of two items. To apply their approach in our context, we consider each topic of interest as an item and the constituent Wikipedia entities of a topic as its content. Then, given $TP^1(u), \dots, TP^L(u)$, we predict $\hat{TP}^{L+1}(u) = (\hat{f}_u^{L+1}(z_1), \dots, \hat{f}_u^{L+1}(z_K))$ over \mathbb{Z}^{L+1} , as follows:

$$\hat{f}_u^{L+1}(z_i) = \frac{1}{L \times K} \sum_{i=1}^L \sum_{j=1}^K f_u^i(z_j) \times S(z_i, z_j) \tag{11}$$

where $S(z_i, z_j)$ denotes the similarity of two topics calculated by the cosine similarity of their respective entity weight distribution vectors.

ACMF This method, which is proposed by Yu et al. (2017), is a hybrid approach that incorporates item-attribute information (item content) into the matrix factorization model to cope with the cold item problem. Specifically, ACMF constrains the baseline matrix factorization framework (MF) with an item relationship regularization term in order to keep two item latent feature vectors relatively close if they are similar in terms of their corresponding attributes. In our experiments, the items are the topics of all time intervals, i.e. $\mathbb{Z} = \bigcup_{1 \leq t \leq L+1} \mathbb{Z}^t$. Accordingly,

the item relationship regularization term is adopted as follows:

$$\frac{\beta}{2} \sum_{i=1}^{|\mathbb{Z}|} \sum_{j=1}^{|\mathbb{Z}|} S(z_i, z_j) \|q_i - q_j\|_F^2 \tag{12}$$

where β is the regularization parameter to control the effect of the item (topic)-attribute information, $S(z_i, z_j)$ is the cosine similarity between the respective entity weight distribution vectors of topics z_i and $z_j \in \mathbb{Z}$. Further, q is the topic latent feature vector, and $\|\cdot\|_F^2$ is the Frobenius norm. We implemented ACMF as an extension to the Librec⁵ implementation of PMF (Salakhutdinov and Mnih 2007) with its default parameter settings (the dimension of the latent feature vector is set to 10, user and item regularization parameters $\lambda_1 = \lambda_2 = 0.1$, item relationship regularization parameter $\beta = 0.2$, and learning rate $\eta = 0.01$).

TLTUP The method which is proposed by Vu et al. (2015) is a content-based recommender system that models the temporal profile of each user based on the latent features learned from her documents of interest. To apply their approach in our context, we consider each topic of interest as a document and the constituent Wikipedia entities of a topic as its content. By applying LDA over all documents, which are the topics of all time intervals, i.e. $\mathbb{Z} = \bigcup_{1 \leq t \leq L+1} \mathbb{Z}^t$, LDA produces the $|\mathbb{Z}|$ topic-feature distributions, where each topic-feature distribution corresponds to a topic $z \in \mathbb{Z}$. Then to build the temporal latent feature profile of each user u (i.e. user-feature distribution), given $TP^1(u), \dots, TP^L(u)$ that represent the topic profiles of user u in L historical time intervals, the probability of a feature f given u is defined as a mixture of probabilities of f given her topics of interest as follows:

$$p(f|u) = \frac{1}{\gamma} \sum_{t=1}^L \sum_{j=1}^K \alpha^{L-t} \times f_u^t(z_j) \times p(f|z_j) \tag{13}$$

where $\gamma = \sum_{t=1}^L \sum_{j=1}^K \alpha^{L-t}$ is a normalisation factor; α denotes the decay rate and $p(f|z)$ represents the probability of feature f given topic z .

Finally, for a given user u , in order to predict $\hat{TP}^{L+1}(u) = (\hat{f}_u^{L+1}(z_1), \dots, \hat{f}_u^{L+1}(z_K))$ over \mathbb{Z}^{L+1} , we calculate $\hat{f}_u^{L+1}(z)$ by measuring the cosine similarity of the feature distribution of topic $z \in \mathbb{Z}$ and feature distribution of user u .

Naive This method is a simple variant of our proposed approach in which for each user u , only those categories that are directly associated with the constituent entities of the user’s topics of interest are considered as categories of interest. Simply put, the hierarchical structure of Wikipedia category hierarchy is not used to propagate the user’s interests. In this method, given $TP^1(u), \dots, TP^L(u)$, we calculate the degree of interest of user u to category $c \in \mathbb{C}$ over L historical time intervals, denoted $g_u^L(c)$, as follows:

$$g_u^L(c) = \sum_{t=1}^L \sum_{i=1}^K f_u^t(z_i) \times \Phi(z_i, c) \tag{14}$$

where $\Phi(z, c)$ denotes the degree of relatedness of topic z to category $c \in \mathbb{C}$ and calculated based on the Embedding mapping function. Then, we predict $\hat{TP}^{L+1}(u)$ based on the topic profile prediction method described in Sect. 4.2.2. We consider this version of our proposed approach as one of our comparison methods to study the effect of considering hierarchical structure of Wikipedia categories in our user modeling on the performance of future interest prediction.

⁵ www.librec.net.

Table 2 The performance of the six model variants in terms of nDCG. The number of topics $K = 20$

		Spreading Function		
		Bell log	Specificity	Priority
Mapping function	Attribution	0.406	0.428	0.459
	Embedding	0.411	0.431	0.477

Table 3 The performance of the six model variants in terms of MAP. The number of topics $K = 20$

		Spreading Function		
		Bell log	Specificity	Priority
Mapping function	Attribution	0.186	0.208	0.241
	Embedding	0.191	0.222	0.261

Hierarchical This method is our proposed approach which uses the hierarchical category profile of users to predict their future interests as described in Sect. 4.2.

As different possible alternative methods are introduced in Sect. 4.2.1 to address the two main functions in our proposed user modeling strategy, i.e. mapping function and spreading function, in the following section, we first conduct an experiment to determine which combination of methods are more influential in accurately predicting the future interests of users on Twitter. Then, in Sect. 5.5, our proposed approach is compared with the baseline methods. Finally, a qualitative analysis provided in Sect. 5.6.

5.4 Analysis of the proposed future interest prediction approach

In this section, we conduct an experiment to explore how and to what extent leveraging different methods for two main functions of our user modeling strategy, i.e. mapping function and spreading function, facilitate the prediction of user future interests on Twitter. To do so, we define and analyze different variants of our prediction model by varying its two main functions. In Sect. 4.2.1, two alternatives, i.e. *Attribution* and *Embedding*, are introduced as mapping functions and three methods, i.e. *Bell-Log*, *Specificity* and *Priority*, are provided as spreading functions. By selecting and combining these different alternatives, we obtain six variants that we will systematically compare in this section.

Tables 2 and 3 provide information about the performance of the six aforementioned model variants in terms of nDCG and MAP, respectively. Based on the results of our analysis in Sect. 5.5, the experimental results presented in this section are obtained with setting the number of topics to 20 and the length of each time interval to 1 week.

The Attribution mapping function maps a given topic to Wikipedia categories that are directly associated with the constituent entities of the topic. However, the Embedding method utilizes an embedding model that includes a distributed representation for Wikipedia entities and categories for this purpose. By comparing the models that leverage Attribution method as their mapping function with those that use the Embedding method, it can be inferred that, by applying all the three selected spreading functions, using the Embedding method outperforms the Attribution method in terms of both nDCG and MAP. This means that using the Embedding model leads to more accurate categories for user topics of interest.

Table 4 Quality of prediction results in terms of nDCG

Method	K = 20	K = 30	K = 40
Hierarchical	0.477 *, ^, ", +	0.478 *, ^, ", +	0.480 *, ^, ", +
Naive	0.423	0.236	0.302
SCRS	0.337	0.215	0.294
ACMF	0.411	0.250	0.362
TLTUP	0.415	0.292	0.381

Based on a t-test, *, ^, " and + indicate a significant difference over Naive, SCRS, ACMF and TLTUP respectively

Table 5 Quality of prediction results in terms of MAP

Method	K=20	K=30	K=40
Hierarchical	0.261 *, ^, ", +	0.336 *, ^, ", +	0.343 *, ^, ", +
Naive	0.233	0.111	0.157
SCRS	0.203	0.109	0.141
ACMF	0.237	0.119	0.170
TLTUP	0.213	0.129	0.181

As another observation, regardless of which mapping function is used to find the related categories of topics, based on the results reported in Tables 2 and 3, it is evident that the Priority spreading function outperforms the other two spreading functions, i.e. Bell-Log and Specificity, in terms of both nDCG and MAP. Based on our observation, given the Priority approach considers the relatedness between a category and its super category to spread the activation values, during the propagation step, some branches of the hierarchy which are not highly related to the category of interest of user are pruned and are not included in the category profile of the user and consequently leading to better performance compared to the methods that do not consider the relatedness between categories to spread the activation values.

In summary, as highlighted in Tables 2 and 3 (the bold values), among different variants analyzed in this section, our proposed future interest prediction model shows the best performance when we employ the *Embedding* method as the mapping function and the *Priority* method as the spreading function. Thus we use this variant as our Hierarchical method to compare with the baselines in the next section.

5.5 Comparison with baseline methods

Our goal in this section is to compare our proposed method with the baseline methods. Tables 4 and 5 report the quality of the prediction results of the proposed model and other competitors in terms of nDCG and MAP respectively, by setting the number of topics to 20, 30 and 40. For all the comparison methods the length of each time interval is set to 1 week which is the optimal time interval for all the methods based on our experiments in Sect. 5.5.1.

Based on the results reported in Tables 4 and 5, we can observe that our proposed approach, i.e. Hierarchical, outperforms all other comparison methods in terms of both nDCG and MAP and by setting the number of topics to 20, 30 and 40. This observation confirms that utilizing Wikipedia category hierarchy can lead to improved quality of user interest prediction with regard to new topics of the future. We also tested the statistical significance of the observed differences between Hierarchical method, and each of the other

comparison methods by performing a paired t-test with 95% confidence level. As depicted in Tables 4 and 5, the improvement of our approach is statistically significant over the other comparison methods in terms of both MAP and nDCG.

By comparing Naive and Hierarchical variants of our proposed approach, one can observe that the Hierarchical method provides better results. Both methods model users' high-level interests over Wikipedia categories. The difference is that, in the Naive method, only those categories that are directly associated with the constituent entities of the user's topics of interest are considered as categories of interest. However, in the Hierarchical method, broadly related categories of user interests are also considered by applying a spreading function over the hierarchy of the Wikipedia category structure. This means that by considering hierarchical structure of Wikipedia categories, we can model high-level interests of users more accurately which consequently leads to improving the results of future interest prediction. However, if we compare the Naive method with the results of the six variants of Hierarchical method reported in Tables 2 and 3, one can observe that the Naive method outperforms some variants of the Hierarchical method. For example, the Naive method provides better results compared to Hierarchical models that leverage Bell-Log method as their spreading activation function in terms of both nDCG and MAP. This means that only considering hierarchical structure of Wikipedia does not lead to better results and inappropriate spreading of scores within the higher levels of the category hierarchy may lead to a poor performance.

The SCRS method is a state of the art work that uses item content to solve the cold item problem. In this method, each topic of interest is considered as an item and its constituent semantic entities are the item content. The fundamental idea of this method is that while topics are changing over time, user's interests follow a consistent pattern. Therefore, this work *realistically* assume that a user will be interested in a new topic if it is similar to her topics of interest in the past. It relies on the pairwise content similarity between the new emerging topics of the future and the past user topics of interest. Two variants of our proposed approach, i.e. Naive and Hierarchical, follows the same intuition, however, we capture this consistent user behavior over time, by generalizing the topic-based representation of user interests into a category-based representation utilizing the Wikipedia categories. Based on the results reported in Tables 4 and 5, we can observe that the two variants of our proposed approach outperform SCRS method in terms of both nDCG and MAP when the number of topics is set to 20, 30 and 40. This observation confirms that utilizing Wikipedia categories enables us to model user's high-level interests more accurately and consequently can lead to improved quality of user interest prediction with regard to new topics in the future.

Now, among the baselines and as shown in Tables 4 and 5, ACMF that is a hybrid recommender system that combines collaborative filtering and topic content, can achieve more accurate results in terms of nDCG and MAP in comparison with SCRS, which is solely based on topic content. This could indicate that incorporating user interests of other users can have impact on the accuracy of user interest predictions. Further, the TLTUP method that incorporates dynamics of user interests by leveraging an exponential decay function to discount the weight of user's interests over time, outperforms the SCRS method in which the interests are not discounted by a decay function. It can be concluded that considering the time decay of user interests improves the quality of user modeling in the context of future interest prediction. This observation is in line with the results reported in Piao and Breslin (2016) that investigate the impact of considering different interest decay functions on the accuracy of user modeling in social networks. Based on these observations, it seems promising to investigate collaborative

Table 6 Effect of the length of time interval in terms of nDCG

Method	1 day	1 week	2 weeks	1 month
Hierarchical	0.292 ^, "	0.477 *, ^, ", +	0.402 *, ^, ", +	0.368 *, ^, ", +
Naive	0.267	0.423	0.343	0.275
SCRS	0.103	0.337	0.248	0.192
ACMF	0.174	0.411	0.307	0.252
TLTUP	0.343	0.415	0.381	0.33

Based on a t-test, *, ^, " and + indicate a significant difference over Naive, SCRS, ACMF and TLTUP respectively

Table 7 Effect of the length of time interval in terms of MAP

Method	1 day	1 week	2 weeks	1 month
Hierarchical	0.109 *, ^, "	0.261 *, ^, ", +	0.215 *, ^, ", +	0.213 *, ^, ", +
Naive	0.098	0.233	0.199	0.189
SCRS	0.054	0.203	0.166	0.137
ACMF	0.072	0.237	0.191	0.164
TLTUP	0.167	0.213	0.191	0.172

extensions of our proposed approach and incorporating interest decay function in our user modeling approach as future work.

As illustrated in Tables 4 and 5, the conclusions drawn from the results are similar for different number of topics in most of cases. Therefore, since in all of the baselines, the best performance is achieved by setting the number of topics to $K = 20$, we only report the experimental results obtained by setting $K = 20$ for the rest of our experiments.

5.5.1 Effect of time interval length

In this section, we study the effect of time interval length on the quality of future interest prediction. The length of time interval controls the time granularity of temporal modeling of historical user activities. We perform the evaluations for different lengths of time interval: 1 day, 1 week, 2 weeks and 1 month. A larger length of time interval indicates that the prediction results will be less time-sensitive. The results in terms of nDCG and MAP are reported in Tables 6 and 7 (in each column, the bold value highlights the best result).

Based on the results, it can be observed that as the length of the time interval increases from 1 day to 1 week, the nDCG and MAP values of all methods increase. One possible reason for this early increase in the quality of prediction results is that increasing the length of each time interval makes the data in a time interval richer and consequently contributes to improved topic discovery and user interest detection. Later on, as the length of the time interval becomes larger than 1 week, the reduced temporal influence leads to the decrease in MAP and nDCG values. All methods achieve their best performance when the length of time interval is set to 1 week (7 days) for partitioning the dataset. As another observation, it can be seen that by setting the length of time

Fig. 6 Effect of the number of recent historical time intervals in terms of nDCG

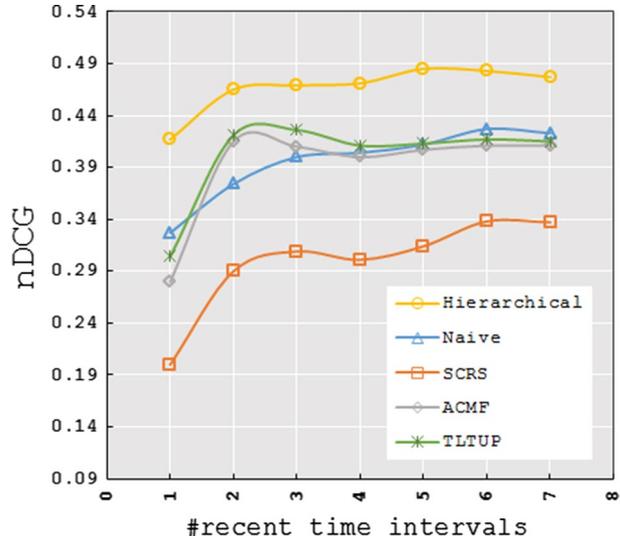
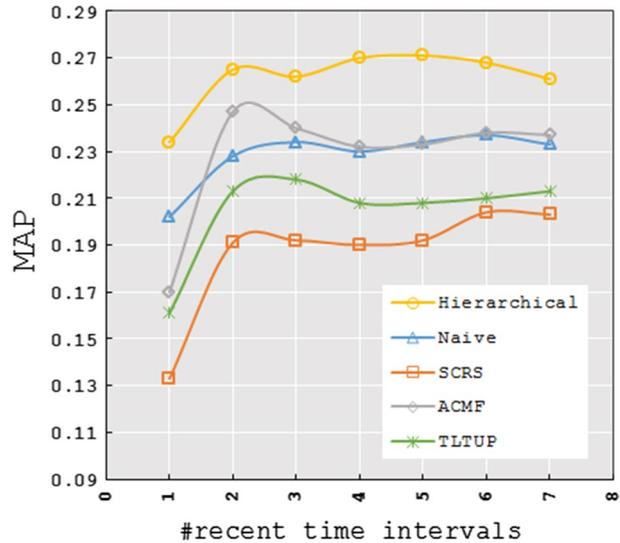


Fig. 7 Effect of the number of recent historical time intervals in terms of MAP



intervals to 1 week, 2 weeks and 1 month, our proposed method, i.e. the Hierarchical method, significantly outperforms all other baselines.

5.5.2 Effect of the number of recent historical time intervals

In the experiments, in order to predict the future topic interest profile of a user u in the testing time interval $L + 1$, we use her historical topic interest profiles in all the consecutive time intervals before the testing time interval in our dataset, i.e. $TP^1(u), \dots, TP^L(u)$. As mentioned earlier, we set the length of each time interval to 1 week for all the comparison

Table 8 Effect of topic activity in terms of nDCG

Method	Cold topics	Semi-active topics	Active topics	Highly-active topics
Hierarchical	0.17 *, ^, ", +	0.21 *, ^, ", +	0.279 *, ^, +	0.272
Naive	0.117	0.153	0.229	0.292
SCRS	0.029	0.043	0.213	0.278
ACMF	0.056	0.07	0.285	0.298
TLTUP	0.09	0.105	0.251	0.285

Based on a t-test, *, ^, " and + indicate a significant difference over Naive, SCRS, ACMF and TLTUP respectively

methods. It is the optimal time interval for all the methods based on our experiments in Sect. 5.5.1. In this section, we analyze the effect of the number of recent historical time intervals on the accuracy of future interest prediction. To do so, we gradually increase the number of recent time intervals from 1 to 7 weeks and evaluate the quality of prediction results. The results in terms of nDCG and MAP are illustrated in Figs. 6 and 7, respectively.

Based on the results, the quality of predicting future interests in terms of both MAP and nDCG significantly increases when the number of training time intervals increases from 1 week to 2 weeks. One possible reason for this improvement is that considering only 1 week of historical data cannot capture the changes of users’ interests and does not include adequate data from user activities. Thus, by increasing the number of recent time intervals from 1 to 2, the models better capture user interests.

As another observation, when the number of training time intervals is larger than 2 weeks, considering more historical user activities does not lead to significant increase or decrease in the quality of predicted future interests in terms of both MAP and nDCG values. This is inline with the fact that user interests change over time and their future interests are more similar to their recent interests compared to the past ones. Therefore, it seems that the consideration of only two weeks before the testing time interval is adequate to capture the historical interests of users for the purpose of future interest prediction in our dataset.

5.5.3 Performance on the cold-start problem

Our goal is to predict user interests over certain topics that emerge in the future on Twitter. The rapid change of trending topics on Twitter causes future interest prediction systems to face the common problem of generating predictions for new topics that may have not been observed in the past. This problem is known as the *cold-start problem*, which means the system does not have enough data to generate predictions for a new item recently added to the system.

To investigate the effectiveness of different comparison methods for coping with the cold-start problem, in this experiment, we study the impact of the topic activity level on the performance of the future interest prediction methods. To do so, we first calculate the level of activity of each topic z_i in the testing time interval, i.e. $z_i \in \mathbb{Z}^{L+1}$, based on the topics which have happened in the training time intervals, i.e. $\mathbb{Z}^1, \dots, \mathbb{Z}^L$, as follows:

$$activity(z_i) = \frac{1}{L} \sum_{t=1}^L MAX_{z_j \in \mathbb{Z}^t} S(z_i, z_j) \tag{15}$$

Table 9 Effect of topic activity in terms of MAP

Method	Cold topics	Semi-active topics	Active topics	Highly-active topics
Hierarchical	0.065 *, ^, ", +	0.087 *, ^, ", +	0.16 *, ^, +	0.168
Naive	0.043	0.059	0.128	0.192
SCRS	0.012	0.018	0.14	0.193
ACMF	0.021	0.027	0.164	0.201
TLTUP	0.048	0.049	0.135	0.165

Table 10 Top-3 topics of two sample users, u_{87} and u_{52} , in three consecutive time intervals before the testing time interval of our Twitter dataset. t_1 : Dec. 6, 2010 to Dec. 13, 2010, t_2 : Dec. 13, 2010 to Dec. 20, 2010 and t_3 : Dec. 20, 2010 to Dec. 27, 2010

Time	u_{87}	u_{52}
t_1	{(z_5 : 0.249), (z_4 : 0.245), (z_6 : 0.087)}	{(z_2 : 0.217), (z_1 : 0.173), (z_3 : 0.130)}
t_2	{(z_9 : 0.344), (z_7 : 0.163), (z_{11} : 0.125)}	{(z_7 : 0.227), (z_8 : 0.188), (z_{10} : 0.089)}
t_3	{(z_{12} : 0.206), (z_{16} : 0.203), (z_{15} : 0.157)}	{(z_{13} , 0.225), (z_{12} , 0.140), (z_{14} , 0.112)}

where $S(z_i, z_j)$ denotes the similarity of two topics z_i and z_j calculated by the cosine similarity of their respective entity weight distribution vectors.

Then, we sort the topics based on their level of activity in ascending order and partition the topics into 4 equal-size groups, namely *cold*, *semi-active*, *active* and *highly-active* topics. Finally, to evaluate the quality of the prediction results in a given group of topics (e.g., cold topics), for each user, we only consider those topics of her ground truth interests which belong to that group of topics to form her new ground truth interests.

Tables 8 and 9 give information about the experimental results in terms of nDCG and MAP, respectively. As reported in these tables, for the first two topic groups, i.e. cold and semi-active topics, our proposed approaches, i.e. Hierarchical and Naive, which utilize user category profile to predict future interests of users, significantly outperform the other three comparison methods in terms of both nDCG and MAP. This observation indicates that our proposed approach can cope with the cold-start problem more effectively than the other comparison methods. We argue that the main reason for the improvement is the consideration of the category of user interests to capture the high-level interests of users instead of only relying on the content of topics. Furthermore, among two variants of our proposed approach, the Hierarchical method performs significantly better than the Naive variant, which means considering the hierarchical structure of Wikipedia categories can have a positive impact for cold and semi-active topics.

Based on the results, we can also observe that as the level of topic activity increases, the improvement of our proposed approach over other comparison methods reduces. For active topics, the Hierarchical approach still significantly outperforms Naive, SCRS and TLTUP approaches and although the nDCG and MAP value of ACMF approach is better than our proposed approach, this difference is not significant. However, for highly-active topics, as highlighted in the tables (the bold values), one can observe that the

ACMF approach, which is a hybrid approach based on both collaborative filtering and similarity between topics, significantly outperforms the Hierarchical method.

5.6 Qualitative analysis

In this section, we analyze two sample user profiles to enable a better interpretation of our proposed approach. Based on the results of our quantitative analysis in Sect. 5.5, in this section, we set the number of topics to 20 and the length of each time interval to 1 week. Table 10 shows the topic profile of user u_{87} and user u_{52} in three consecutive weeks before the testing time interval in our Twitter dataset (i.e. the last week in December 2010), which are extracted by Twitter-LDA as described in Sect. 4.1. Further, the information about top-5 constituent entities of the topics are illustrated in Table 11. In Table 10, for each user u in each time interval t , we only show the top-3 topics with the highest probability in her topic profile $TP^t(u)$. The weight of each topic indicates users' preference for the topic. For example ($z_5, 0.249$) for user u_{87} in time interval t_1 means the degree of interest of user u_{87} to topic z_5 is 0.249. Note that we only choose the top-3 topics for demonstration purposes, and thus the sum of the weights for topics of interest of each user in each time interval is not equal to one.

Based on Tables 10 and 11, by comparing the interests of users u_{87} and u_{52} in each time interval, it can be seen that their dominating topics of interest are different and there is only one overlapping topic for these users in time intervals t_2 and t_3 . Further, if we compare the topics of interest of each user in different time intervals, one can see that the dominating interests of users change over these three adjacent time intervals. For example, the dominating interest of user u_{52} in time interval t_1 is topic z_2 which is related to the event where WikiLeaks supporters hacked a credit card website, which became a trending topic in that time interval. However, subsequently, the user's main interest totally changes to the 2010 New York Jets season, the 41st season in the National Football League, in time interval t_3 .

It should be noted that by manually checking the topics extracted by Twitter-LDA for each time interval, we made some interesting observations. Some topics represent an abstract topic, while others are time-oriented topics that represent a real-world event in that given time interval. The information from some sample topics extracted by Twitter-LDA from our dataset is illustrated in Table 11. For example, topic $z_{12} = \{\text{Health, Human, Happiness, Life, Woman}\}$ that refers to *human health and happiness* is an abstract topic that may be interesting for users in social networks at any given time. However, topic $z_{15} = \{\text{Christmas, Christmas_Eve, Santa_Claus, Gift, Christmas_and_holiday_season}\}$ is a topic that becomes trendy on Twitter late December, due to the new year events. It is worth noting that because our extracted topics include both types of topics, we use topic and event interchangeably in the paper, to refer to the topics of interest to the users.

Our proposed prediction approach is based on the intuition that while user interests might change over time, as demonstrated for our two sample users u_{87} and u_{52} in Table 10, users tend to revolve around some fundamental concepts that can be seen as their mindset or their high-level interests. To model the high-level interests of each user, we utilized Wikipedia categories. The main step for doing so is to map each topic to its related categories. Table 11 (third column) gives information about the top-5 related Wikipedia categories of each topic, which are identified by our proposed embedding method. As we use the

Table 11 Top-5 entities of topics for sample users u_{87} and u_{52} extracted for three time intervals t_1 , t_2 and t_3 and their related Wikipedia categories. t_1 : Dec. 6, 2010 to Dec. 13, 2010, t_2 : Dec. 13, 2010 to Dec. 20, 2010 and t_3 : Dec. 20, 2010 to Dec. 27, 2010

Topic#	Top-5 entities	Wikipedia categories	Time
z_1	Barack_Obama, Republican_Party_(United_States), Tax_cut, Tax, United_States_Senate	United_States_Senate, United_States_House_of_Representatives, Democratic_Party_(United_States), Republican_Party_(United_States), Republican_Party_United_States, United_States_Senators	t_1
z_2	WikiLeaks, MasterCard, Hacker_(computer_security), Julian_Assange, PayPal	WikiLeaks, Hacking_in_the_2010s, News_leaks, Online_organizations, Media_critics	t_1
z_3	Android_(operating_system), iPhone, iPad, Apple_Inc., Mobile_phone	Apple_Inc., mobile_phones, Smartphones, Multi-touch_mobile_phones, 2008_software, Cloud_clients	t_1
z_4	SpaceX, NASA, Falcon_9, Rocket_launch, Dragon_(spacecraft)	SpaceDev_aircraft, SpaceX_payloads_contracted_by_NASA, SpaceX, Space_launch_System, Companies_established_in_2002	t_1
z_5	Elizabeth_Edwards, Cancer, Death, Breast, Don't_ask,_don't_tell	Writers_from_North_Carolina, Deaths_from_cancer_in_North_Carolina, Spouses_of_North_Carolina_politicians, Hawaii_state_case_law, People_from_Raleigh,_North_Carolina	t_1
z_6	Attack_on_Pearl_Harbor, Photography, Photograph, Shark, Pearl_Harbor	Attack_on_Pearl_Harbor, Ports_and_harbors_of_Hawaii, 1941_in_Hawaii, History_of_Midway_Atoll, 1941_in_the_United_States	t_1
z_7	Tackle_(American_football), National_Football_League, Brett_Favre, American_football, New_York_Jets	American_football_leagues_in_the_United_States, Sports_leagues_established_in_1920, National_Football_League, American_football_teams_in_New_Jersey, American_football_teams_in_New_York_City	t_2
z_8	Indonesia, Jakarta, Cuba, YG_(rapper), Poetry	Indonesia, Vietnam, Republics, Cuba, Thailand	t_2
z_9	Woman, Christmas, Murder, Death, Student	Women, Females, Conflict_in_society, Quarter_days, Role_status	t_2
z_{10}	Afghanistan, Barack_Obama, Richard_Holbrooke, Pakistan, War_in_Afghanistan	Pashto-speaking_countries_and_territories, Territories_under_military_occupation, Urdu-speaking_countries_and_territories, Republics, Afghanistan	t_2

Table 11 (continued)

Topic#	Top-5 entities	Wikipedia categories	Time
z ₁₁	IPhone, iPad, Apple Inc., Android (operating system), Application software	Apple_Inc._mobile_phones, Multi-touch_mobile_phones, Companies_based_in_Cupertino, California, Products_introduced_in_2007, Cloud_clients	t ₂
z ₁₂	Health, Human, Happiness, Life, Woman	Pleasure, Feeling, Humans, Core_issues_in_ethics, Axiological_theories	t ₃
z ₁₃	Tackle (American football), National Football League, Los Angeles Lakers, American football, New York Jets	American_football_leagues_in_the_United_States, Sports_leagues_established_in_1920, National_Football_League, American_football_teams_in_Washington, D.C., Organizations_based_in_New_York_City	t ₃
z ₁₄	Barack Obama, Republican Party (United States), United States Senate, United States Congress, President of the United States	United_States_Senate, Republican_Party_United_States_Senators, United_States_House_of_Representatives, Cabinet_of_the_United_States, United_States_Senators_from_Illinois	t ₃
z ₁₅	Christmas, Christmas Eve, Santa Claus, Gift, Christmas and holiday season	Quarter_days, Christmas, Seasonal_traditions, Christmas-linked_holidays, Holiday-related_topics	t ₃
z ₁₆	China, Business, Employment, India, Economy	China, Communist_states, Central_Asian_countries, Republics, Chinese-speaking_countries_and_territories	t ₃

constituent entities of each topic in our mapping function, it is obvious that the quality of the extracted topics directly impacts the quality of extracting related categories.

For example, for topic $z_{15} = \{\text{Christmas, Christmas_Eve, Santa_Claus, Gift, Christmas_and_holiday_season}\}$ that properly represents the topic of *Christmas Eve*, all the extracted categories for this topic, i.e. *Quarter_days, Christmas, Seasonal_traditions, Christmas-linked_holidays* and *Holiday-related_topics* are related to the topic. However, for topic $z_8 = \{\text{Indonesia, Jakarta, Cuba, YG_rapper}, \text{Poetry}\}$, which seems to be a mixed topic that represents more than one event, the extracted categories, i.e. *Indonesia, Jakarta, Cuba, YG_rapper* and *Poetry*, do not seem very informative. To alleviate this problem, our mapping function takes into account the participation score of each entity in each topic. For example in the topic $z_5 = \{\text{Elizabeth_Edwards, Cancer, Death, Breast, Don't_ask, don't_tell}\}$ that refers to the death of Elizabeth Edward because of her breast cancer, the last entity, i.e. *Don't_ask_don't_tell* is not related to this event. However, because its participation score in the topic is very low compared to the participation score of Elizabeth Edward, as one can see in the Table 11, the extracted categories for this topic are not affected by this mistake and they are correctly related to the topic.

To predict user interests over future topics, we use the similarity between the categories of each future topic and category profile of the user which is built based on the related categories of her interests in the past. For example, one trending topic in the last week of December (our testing time interval), extracted by applying twitter-LDA over the tweets published in this week, is the *New Year* topic, which includes *New_Year's_Eve, New_Year's_Day, Happy_New_Year_(song), With_You_(Chris_Brown_song), Happiness* as its top-5 entities. User u_{87} who is interested in the Christmas Eve topic would be interested in the New Year topic, too. However the prediction models that only use the content of topics, cannot predict that user u_{87} would be interested in New Year topic. Because these two topics don't have any overlapping entities. Our mapping function identifies the following categories for the topic of New Year: *New_Year_celebrations, Holiday-related_topics, January_observances, Winter_holidays* and *Annual_events*. Because these two topics, i.e. New Year and Christmas Eve, have a common category, i.e. *holiday_related_topics*, our approach can infer that user u_{87} would also be interested in the New Year topic in the last week of December.

6 Conclusions and future work

In this paper, we address the problem of predicting future interests of users with regard to a set of topics that emerge in the future, which may have not been observed in the past. Our prediction model is based on the intuition that while user interests might change over time in reaction to real world events, they tend to revolve around some fundamental concepts that can be seen as the user's mindset. To capture the temporal behaviour of the topics and user interests on Twitter, we consider discrete time intervals and build topic interest profile of each user in these time intervals as the user's historical topic profiles. To capture the user's overall mindset, we generalize each individual user topic interest profile as we move through time from the oldest to the most recent interval to infer the user category profile using the Wikipedia category structure. Given a user category profile, we predict the degree of interest of the user to each future topic based on the relatedness of each topic

to the inferred category profile. Our experiments illustrate better performance for our proposed approach compared to the state of the art.

There are several directions, which we would like to explore in the future. Given the fact that social networks are highly skewed and imbalanced media, it is interesting to investigate how skewed data will affect the quality of user interest profiles. In this paper and in order to alleviate this issue, we focused only on active users. Studies have reported that many users on social networks can exhibit cold start or free riding behavior (Romero et al. 2011). Therefore, we intend to utilize our recently proposed approach in Zarrinkalam et al. (2018) to identify user interests for those users who are not highly active on the social network either because they are cold start users or free riders, and predict the future interests for all the users, instead of only for active users. As another future work, we intend to include social connections between users in our framework, and extend our model to a hybrid recommender system that combines collaborative filtering, users' social connections, semantic information from knowledge bases and temporal features of users' interests at the same time to model high-level interests of users and consequently predict users' future interests with regard to unobserved topics.

Acknowledgements We are immensely thankful to our colleague at University of New Brunswick, Hossein Fani (<https://sites.google.com/site/hosseinfani/>), whose comments helped improve and clarify this manuscript.

References

- Abdel-Hafez, A., & Xu, Y. (2013). A survey of user modelling in social media websites. *Computer and Information Science*, 6(4), 59–71. <https://doi.org/10.5539/cis.v6n4p59>.
- Abel, F., Gao, Q., Houben, G., & Tao, K. (2011). Analyzing temporal dynamics in twitter profiles for personalized recommendations in the social web. In *Web Science 2011, WebSci '11*, Koblenz, Germany, 2011 (pp. 2:1–2:8). <https://doi.org/10.1145/2527031.2527040>.
- Abel, F., Gao, Q., Houben, G., & Tao, K. (2011). Analyzing user modeling on twitter for personalized news recommendations. In *Proceedings user modeling, adaption and personalization-19th international conference, umap 2011*, Girona, Spain, July 11–15, 2011 (pp. 1–12). https://doi.org/10.1007/978-3-642-22362-4_1.
- Abel, F., Gao, Q., Houben, G., & Tao, K. (2011). Semantic enrichment of twitter posts for user profile construction on the social web. In *Proceedings, Part II, the semantic web: Research and applications—8th extended semantic web conference, ESWC 2011*, Heraklion, Crete, Greece 2011 (pp. 375–389). https://doi.org/10.1007/978-3-642-21064-8_26.
- Akbari, M., Hu, X., Wang, F., & Chua, T. (2017). Wellness representation of users in social media: Towards joint modelling of heterogeneity and temporality. *IEEE Transactions on Knowledge and Data Engineering*, 29(10), 2360–2373. <https://doi.org/10.1109/TKDE.2017.2722411>.
- Asur, S., & Huberman, B. A. (2010). Predicting the future with social media. In *Main conference proceedings on 2010 IEEE/WIC/ACM international conference on web intelligence, WI 2010*, Toronto, Canada, August 31–September 3, 2010 (pp. 492–499). <https://doi.org/10.1109/WI-IAT.2010.63>.
- Bao, H., Li, Q., Liao, S. S., Song, S., & Gao, H. (2013). A new temporal and social pmf-based method to predict users' interests in micro-blogging. *Decision Support Systems*, 55(3), 698–709. <https://doi.org/10.1016/j.dss.2013.02.007>.
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77–84. <https://doi.org/10.1145/2133806.2133826>.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022.
- Bobadilla, J., Ortega, F., Hernando, A., & Gutiérrez, A. (2013). Recommender systems survey. *Knowledge-based Systems*, 46, 109–132. <https://doi.org/10.1016/j.knosys.2013.03.012>.

- Boldi, P., & Monti, C. (2016). Cleansing wikipedia categories using centrality. In *Proceedings of the 25th international conference on World Wide Web, WWW 2016*, Montreal, Canada, April 11–15, 2016, Companion Volume (pp. 969–974). <https://doi.org/10.1145/2872518.2891111>.
- Cheekula, S. K. (2016). Characterizing concepts in taxonomy for entity recommendations. Master's thesis, Wright State University, Dayton
- Cheekula, S. K., Kapanipathi, P., Doran, D., Jain, P., & Sheth, A. P. (2015). Entity recommendations using hierarchical knowledge bases. In *Proceedings of the 4th workshop on knowledge discovery and data mining meets linked open data co-located with 12th extended semantic web conference (ESWC 2015)*, Portoroz, Slovenia, May 31, 2015.
- Chen, J., Nairn, R., Nelson, L., Bernstein, M. S., & Chi, E. H. (2010). Short and tweet: experiments on recommending content from information streams. In *Proceedings of the 28th international conference on human factors in computing systems, CHI 2010*, Atlanta, Georgia, USA, April 10–15, 2010 (pp. 1185–1194). <https://doi.org/10.1145/1753326.1753503>.
- Cheng, X., Yan, X., Lan, Y., & Guo, J. (2014). BTM: Topic modeling over short texts. *IEEE Transactions on Knowledge and Data Engineering*, 26(12), 2928–2941. <https://doi.org/10.1109/TKDE.2014.2313872>.
- Cornolti, M., Ferragina, P., & Ciaramita, M. (2013). A framework for benchmarking entity-annotation systems. In *22nd International World Wide Web Conference, WWW '13*, Rio de Janeiro, Brazil, May 13–17, 2013 (pp. 249–260).
- Ekstrand, M. D., Riedl, J., & Konstan, J. A. (2011). Collaborative filtering recommender systems. *Foundations and Trends in Human-Computer Interaction*, 4(2), 175–243. <https://doi.org/10.1561/1100000009>.
- Fani, H., Zarrinkalam, F., Bagheri, E., & Du, W. (2016). Time-sensitive topic-based communities on twitter. In *Proceedings on advances in artificial intelligence-29th Canadian conference on artificial intelligence, Canadian AI 2016*, Victoria, BC, Canada, May 31–June 3, 2016 (pp. 192–204). https://doi.org/10.1007/978-3-319-34111-8_25.
- Faralli, S., Stilo, G., & Velardi, P. (2017). Automatic acquisition of a taxonomy of microblogs users interests. Web Semantics: Science, Services and Agents on the World Wide Web.
- Ferragina, P., & Scaife, U. (2012). Fast and accurate annotation of short texts with wikipedia pages. *IEEE Software*, 29(1), 70–75. <https://doi.org/10.1109/MS.2011.122>.
- Gao, L., Wu, J., Zhou, C., & Hu, Y. (2017). Collaborative dynamic sparse topic regression with user profile evolution for item recommendation. In *Proceedings of the thirty-first AAAI conference on artificial intelligence*, February 4–9, 2017, San Francisco, California, USA (pp. 1316–1322).
- Huang, J., Peng, M., Wang, H., Cao, J., Gao, W., & Zhang, X. (2017). A probabilistic method for emerging topic tracking in microblog stream. *World Wide Web*, 20(2), 325–350. <https://doi.org/10.1007/s11280-016-0390-4>.
- Kapanipathi, P., Jain, P., Venkatramani, C., & Sheth, A. P. (2014). User interests identification on twitter using a hierarchical knowledge base. In *Proceedings on the semantic web: Trends and challenges-11th international conference, ESWC 2014*, Anissaras, Crete, Greece, May 25–29, 2014 (pp. 99–113). https://doi.org/10.1007/978-3-319-07443-6_8.
- Kapanipathi, P., Orlandi, F., Sheth, A. P., & Passant, A. (2011). Personalized filtering of the twitter stream. In *Proceedings of the second workshop on semantic personalized information management: Retrieval and recommendation 2011*, Bonn, Germany, October 24, 2011 (pp. 6–13).
- Koren, Y. (2010). Collaborative filtering with temporal dynamics. *Communications of the ACM*, 53(4), 89–97. <https://doi.org/10.1145/1721654.1721677>.
- Li, L., Zheng, L., Yang, F., & Li, T. (2014). Modeling and broadening temporal user interest in personalized news recommendation. *Expert Systems with Applications*, 41(7), 3168–3177. <https://doi.org/10.1016/j.eswa.2013.11.020>.
- Li, Y., Zheng, R., Tian, T., Hu, Z., Iyer, R., & Sycara, K. P. (2016). Joint embedding of hierarchical categories and entities for concept categorization and dataless classification. In *Proceedings of the 26th international conference on computational linguistics, COLING 2016*, December 11–16, 2016, Osaka, Japan (pp. 2678–2688).
- Liang, S., Ren, Z., Zhao, Y., Ma, J., Yilmaz, E., & de Rijke, M. (2017). Inferring dynamic user interests in streams of short texts for user clustering. *ACM Transactions on Information Systems (TOIS)*, 36(1), 10:1–10:37.
- Lu, C., Lam, W., & Zhang, Y. (2012). Twitter user modeling and tweets recommendation based on wikipedia concept graph. In *Workshops at the Twenty-Sixth AAAI Conference on Artificial Intelligence*.
- Mehrotra, R., Sanner, S., Buntine, W. L., & Xie, L. (2013). Improving LDA topic models for microblogs via tweet pooling and automatic labeling. In *The 36th International ACM SIGIR conference on research and development in Information Retrieval, SIGIR '13*, Dublin, Ireland - July 28–August 01, 2013 (pp. 889–892). <https://doi.org/10.1145/2484028.2484166>.

- Michelson, M., & Macskassy, S. A. (2010). Discovering users' topics of interest on twitter: a first look. In *Proceedings of the Fourth Workshop on Analytics for Noisy Unstructured Text Data, AND 2010*, Toronto, Ontario, Canada, October 26th, 2010 (in conjunction with CIKM 2010) (pp. 73–80). <https://doi.org/10.1145/1871840.1871852>.
- Mikolov, T., Chen, K., & Corrado, G., Dean, J. (2013). Efficient estimation of word representations in vector space. CoRR arXiv:1301.3781.
- Nakatsuji, M., Fujiwara, Y., Uchiyama, T., & Toda, H. (2012). Collaborative filtering by analyzing dynamic user interests modeled by taxonomy. In *Proceedings, Part I, The Semantic Web-ISWC 2012—11th international semantic web conference*, Boston, MA, USA, November 11–15, 2012 (pp. 361–377). https://doi.org/10.1007/978-3-642-35176-1_23.
- Nguyen, P. T., Tomeo, P., Noia, T. D., & Sciascio, E. D. (2015). Content-based recommendations via dbpedia and freebase: A case study in the music domain. In *Proceedings, Part I the semantic web—ISWC 2015—14th international semantic web conference*, Bethlehem, PA, USA, October 11–15, 2015 (pp. 605–621). https://doi.org/10.1007/978-3-319-25007-6_35.
- Noia, T. D., Mirizzi, R., Ostuni, V. C., Romito, D., & Zanker, M. (2012). Linked open data to support content-based recommender systems. In *I-SEMANTICS 2012—8th international conference on semantic systems, I-SEMANTICS '12*, Graz, Austria, September 5–7, 2012 (pp. 1–8). <https://doi.org/10.1145/2362499.2362501>.
- Ostuni, V. C., Noia, T. D., Sciascio, E. D., & Mirizzi, R. (2013). Top-n recommendations from implicit feedback leveraging linked open data. In *Seventh ACM Conference on Recommender Systems, RecSys '13*, Hong Kong, China, October 12–16, 2013 (pp. 85–92). <https://doi.org/10.1145/2507157.2507172>.
- Passant, A. (2010). Dbrec-music recommendations using dbpedia. In *The semantic Web - ISWC 2010—9th international semantic web conference, ISWC 2010*, Shanghai, China, November 7–11, 2010, Revised Selected Papers, Part II (pp. 209–224). https://doi.org/10.1007/978-3-642-17749-1_14.
- Piao, G., & Breslin, J. G. (2016). Exploring dynamics and semantics of user interests for user modeling on twitter for link recommendations. In *Proceedings of the 12th international conference on semantic systems, SEMANTICS 2016*, Leipzig, Germany, September 12–15, 2016 (pp. 81–88). <https://doi.org/10.1145/2993318.2993332>.
- Piao, G., & Breslin, J. G. (2017). Inferring user interests for passive users on twitter by leveraging follower biographies. In *Proceedings advances in information retrieval—39th European conference on IR research, ECIR 2017*, Aberdeen, UK, April 8–13, 2017 (pp. 122–133). https://doi.org/10.1007/978-3-319-56608-5_10.
- Ponzetto, S. P., & Strube, M. (2007). Deriving a large-scale taxonomy from wikipedia. In *Proceedings of the twenty-second AAAI conference on artificial intelligence*, July 22–26, 2007, Vancouver, British Columbia, Canada (pp. 1440–1445).
- Ramage, D., Dumais, S. T., & Liebling, D. J. (2010). Characterizing microblogs with topic models. In *Proceedings of the Fourth international conference on weblogs and social media, ICWSM 2010*, Washington, DC, USA, May 23–26, 2010.
- Romero, D. M., Galuba, W., Asur, S., & Huberman, B. A. (2011). Influence and passivity in social media. In *Proceedings, Part III Machine learning and knowledge discovery in databases—European Conference, ECML PKDD 2011*, Athens, Greece, September 5–9, 2011 (pp. 18–33). https://doi.org/10.1007/978-3-642-23808-6_2.
- Salakhutdinov, R., & Mnih, A. (2007). Probabilistic matrix factorization. In *Advances in neural information processing systems 20, proceedings of the twenty-first annual conference on neural information processing systems*, Vancouver, British Columbia, Canada, December 3–6, 2007 (pp. 1257–1264).
- Sang, J., Lu, D., & Xu, C. (2015). A probabilistic framework for temporal user modeling on microblogs. In *Proceedings of the 24th ACM international conference on information and knowledge management, CIKM 2015*, Melbourne, VIC, Australia, October 19–23, 2015 (pp. 961–970). <https://doi.org/10.1145/2806416.2806470>.
- Siehndel, P., & Kawase, R. (2012). TwikiMe!—user profiles that make sense. In *Proceedings of the ISWC 2012 posters & demonstrations track*, Boston, USA, November 11–15, 2012.
- Song, X., Nie, L., Zhang, L., Liu, M., & Chua, T. (2015). Interest inference via structure-constrained multi-source multi-task learning. In *IJCAI* (pp. 2371–2377). AAAI Press.
- Song, Y., Elkahky, A. M., & He, X. (2016). Multi-rate deep learning for temporal recommendation. In *Proceedings of the 39th international ACM SIGIR conference on research and development in information retrieval, SIGIR 2016*, Pisa, Italy, July 17–21, 2016 (pp. 909–912). <https://doi.org/10.1145/2911451.2914726>.
- Sriram, B., Fuhry, D., Demir, E., Ferhatosmanoglu, H., & Demirbas, M. (2010). Short text classification in twitter to improve information filtering. In *Proceeding of the 33rd international ACM SIGIR*

- conference on research and development in information retrieval, SIGIR 2010, Geneva, Switzerland, July 19–23, 2010 (pp. 841–842). <https://doi.org/10.1145/1835449.1835643>.
- Varga, A., Basave, A. E. C., Rowe, M., Ciravegna, F., & He, Y. (2014). Linked knowledge sources for topic classification of microposts: A semantic graph-based approach. *Web Semantics: Science, Services and Agents on the World Wide Web*, 26, 36–57. <https://doi.org/10.1016/j.websem.2014.04.001>.
- Vu, T. T., Willis, A., Tran, S. N., & Song, D. (2015). Temporal latent topic user profiles for search personalisation. *ECIR, Lecture Notes in Computer Science*, 9022, 605–616.
- Weng, J., Lim, E., Jiang, J., & He, Q. (2010). Twitterrank: finding topic-sensitive influential twitterers. In *Proceedings of the third international conference on web search and web data mining, WSDM 2010*, New York, NY, USA, February 4–6, 2010 (pp. 261–270). <https://doi.org/10.1145/1718487.1718520>.
- Xiang, L., Yuan, Q., Zhao, S., Chen, L., Zhang, X., Yang, Q., & Sun, J. (2010). Temporal recommendation on graphs via long- and short-term preference fusion. In *Proceedings of the 16th ACM SIGKDD international conference on knowledge discovery and data mining*, Washington, DC, USA, July 25–28, 2010 (pp. 723–732). <https://doi.org/10.1145/1835804.1835896>.
- Xu, Z., Lu, R., Xiang, L., & Yang, Q. (2011). Discovering user interest on twitter with a modified author-topic model. In *Proceedings of the 2011 IEEE/WIC/ACM international conference on web intelligence, WI 2011*, Campus Scientifique de la Doua, Lyon, France, August 22–27, 2011 (pp. 422–429). <https://doi.org/10.1109/WI-IAT.2011.47>.
- Yang, L., Sun, T., Zhang, M., & Mei, Q. (2012). We know what @you #tag: does the dual role affect hashtag adoption? In *Proceedings of the 21st World Wide Web Conference 2012, WWW 2012*, Lyon, France, April 16–20, 2012 (pp. 261–270). <https://doi.org/10.1145/2187836.2187872>.
- Yin, H., Cui, B., Chen, L., Hu, Z., & Zhou, X. (2015). Dynamic user modeling in social media systems. *ACM Transactions on Information Systems (TOIS)*, 33(3), 10:1–10:44. <https://doi.org/10.1145/2699670>.
- Yu, Y., Wang, C., Wang, H., & Gao, Y. (2017). Attributes coupling based matrix factorization for item recommendation. *Applied Intelligence*, 46(3), 521–533. <https://doi.org/10.1007/s10489-016-0841-8>.
- Zarrinkalam, F., Fani, H., Bagheri, E., & Kahani, M. (2016). Inferring implicit topical interests on twitter. In *Proceedings advances in information retrieval–38th European conference on IR research, ECIR 2016*, Padua, Italy, March 20–23, 2016 (pp. 479–491). https://doi.org/10.1007/978-3-319-30671-1_35.
- Zarrinkalam, F., Fani, H., Bagheri, E., & Kahani, M. (2017). Predicting users future interests on twitter. In *In 39th European Conference on Information Retrieval (ECIR2017)*.
- Zarrinkalam, F., Fani, H., Bagheri, E., Kahani, M., & Du, W. (2015). Semantics-enabled user interest detection from twitter. In *IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology, WI-IAT 2015*, Singapore, December 6–9, 2015, Volume I (pp. 469–476). <https://doi.org/10.1109/WI-IAT.2015.182>.
- Zarrinkalam, F., Kahani, M., & Bagheri, E. (2018). Mining user interests over active topics on social networks. *Information Processing & Management*, 54(2), 339–357. <https://doi.org/10.1016/j.ipm.2017.12.003>.
- Zhang, F., Yuan, N. J., Lian, D., Xie, X., & Ma, W. (2016). Collaborative knowledge base embedding for recommender systems. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, San Francisco, CA, USA, August 13–17, 2016 (pp. 353–362). <https://doi.org/10.1145/2939672.2939673>.
- Zhao, W. X., Jiang, J., Weng, J., He, J., Lim, E., Yan, H., & Li, X. (2011). Comparing twitter and traditional media using topic models. In *Proceedings advances in information retrieval–33rd European conference on IR research, ECIR 2011*, Dublin, Ireland, April 18–21, 2011 (pp. 338–349). https://doi.org/10.1007/978-3-642-20161-5_34.