

Analyzing User Modeling on Twitter for Personalized News Recommendations

Fabian Abel, Qi Gao, Geert-Jan Houben, Ke Tao

Web Information Systems, Delft University of Technology
{f.abel,q.gao,g.j.p.m.houben,k.tao}@tudelft.nl

Abstract. How can micro-blogging activities on Twitter be leveraged for user modeling and personalization? In this paper we investigate this question and introduce a framework for user modeling on Twitter which enriches the semantics of Twitter messages (tweets) and identifies topics and entities (e.g. persons, events, products) mentioned in tweets. We analyze how strategies for constructing hashtag-based, entity-based or topic-based user profiles benefit from semantic enrichment and explore the temporal dynamics of those profiles. We further measure and compare the performance of the user modeling strategies in context of a personalized news recommendation system. Our results reveal how semantic enrichment enhances the variety and quality of the generated user profiles. Further, we see how the different user modeling strategies impact personalization and discover that the consideration of temporal profile patterns can improve recommendation quality.

Key words: user modeling, twitter, semantic enrichment, personalized news recommendation, social web, analysis

1 Introduction

With more than 190 million users and more than 65 million postings per day, Twitter is today the most prominent micro-blogging service available on the Web¹. People publish short messages (tweets) about their everyday activities on Twitter and lately researchers investigate feasibility of applications such as trend analysis [1] or Twitter-based early warning systems [2]. Most research initiatives study network structures and properties of the Twitter network [3–5]. Yet, little research has been done on understanding the semantics of individual Twitter activities and inferring user interests from these activities. As tweets are limited to 140 characters, making sense of individual tweets and exploiting tweets for user modeling are non-trivial problems.

In this paper we study how to leverage Twitter activities for user modeling and evaluate the quality of user models in the context of recommending news articles. We develop a framework that enriches the semantics of individual Twitter activities and allows for the construction of different types of semantic user profiles. The characteristics of these user profiles are influenced by different design dimensions and design alternatives. To better understand how those

¹ <http://techcrunch.com/2010/06/08/twitter-190-million-users/>

factors impact the characteristics and quality of the resulting user profiles, we conduct an in-depth analysis on a large Twitter dataset of more than 2 million tweets and answer research questions such as the following: how does the semantic enrichment impact the characteristics and quality of Twitter-based profiles (see Section 4.2)? How do (different types of) profiles evolve over time? Are there any characteristic temporal patterns (see Section 4.3)? How do the different user modeling strategies impact personalization (personalized news article recommendations) and does the consideration of temporal patterns improve the accuracy of the recommendations (see Section 5)?

Before studying the above research questions in Section 4-5, we will summarize related work in Section 2 and introduce the design dimensions of Twitter-based user modeling as well as our Twitter user modeling framework in Section 3.

2 Related Work

With the launch of Twitter in 2007, micro-blogging became highly popular and researchers started to investigate Twitter’s information propagation patterns [3] or analyzed structures of the Twitter network to identify influential users [4, 5]. Dong et al. [6] exploit Twitter to detect and rank fresh URLs that have possibly not been indexed by Web search engines yet. Lately, Chen et al. conducted a study on recommending URLs posted in Twitter messages and compare strategies for selecting and ranking URLs by exploiting the social network of a user as well as the general popularity of the URLs in Twitter [7]. Chen et al. do not investigate user modeling in detail, but represent Twitter messages of a user by means of a bag of words. In this paper we go beyond such representations and analyze different types of profiles like entity-based or hashtag-based profiles.

Laniado and Mika introduce metrics to describe the characteristics of hashtags – keywords starting with “#” – such as frequency, specificity or stability over time [8]. Huang et al. further characterize the temporal dynamics of hashtags via statistical measures such as standard deviation and discover that some hashtags are used widely for a few days but then disappear quickly [9]. Recent research on collaborative filtering showed that the consideration of such temporal dynamics impacts recommendation quality significantly [10]. However, the impact of temporal characteristics of Twitter-based user profiles on recommendation performance has not been researched yet.

Neither hashtag-based nor bag-of-words representation explicitly specify the semantics of tweets. To better understand the semantics of Twitter messages published during scientific conferences, Rowe et al. [11] map tweets to conference talks and exploit metadata of the corresponding research papers to enrich the semantics of tweets. Rowe et al. mention user profiling as one of the applications that might benefit from such semantics, but do not further investigate user modeling on Twitter. In this paper we close this gap and present the first large-scale study on user modeling based on Twitter activities and moreover explore how different user models impact the accuracy of recommending news articles.

design dimension	design alternatives (discussed in this paper)
profile type	(i) hashtag-based, (ii) topic-based or (iii) entity-based
enrichment	(i) tweet-only-based enrichment or (ii) linkage and exploitation of external news articles (propagating entities/topics)
temporal constraints	(i) specific time period(s), (ii) temporal patterns (<i>weekend</i> , <i>night</i> , etc.) or (iii) no constraints

Table 1. Design space of Twitter-based user modeling strategies.

3 Twitter-based User Modeling

The user modeling strategies proposed and discussed in this paper vary in three design dimensions: (i) the type of profiles created by the strategies, (ii) the data sources exploited to further enrich the Twitter-based profiles and (iii) temporal constraints that are considered when constructing the profiles (see Table 1). The generic model for profiles representing users is specified in Definition 1.

Definition 1 (User Profile). *The profile of a user $u \in U$ is a set of weighted concepts where with respect to the given user u for a concept $c \in C$ its weight $w(u, c)$ is computed by a certain function w .*

$$P(u) = \{(c, w(u, c)) | c \in C, u \in U\}$$

Here, C and U denote the set of concepts and users respectively.

In particular, following Table 1 we analyze three *types* of profiles that differ with respect to the type of concepts C : entity-, topic- and hashtag-based profiles – denoted by $P_E(u)$, $P_T(u)$ and $P_H(u)$ respectively. We apply occurrence frequency as weighting scheme $w(u, c)$, which means that the weight of a concept is determined by the number of Twitter activities in which user u refers to concept c . For example, in a hashtag-based profile $w(u, \#technology) = 5$ means that u published five Twitter messages that mention “#technology”. We further normalize user profiles so that the sum of all weights in a profile is equal to 1: $\sum_{c_i \in C} w(u, c_i) = 1$. With $\mathbf{p}(u)$ we refer to $P(u)$ in its vector space model representation, where the value of the i -th dimension refers to $w(u, c_i)$.

The user modeling strategies we analyze in this paper exploit Twitter messages posted by a user u to construct the corresponding profile $P(u)$. When constructing entity- and topic-based user profiles, we also investigate the impact of further *enrichment* based on the exploitation of external data sources (see Table 1). In particular, we allow for enrichment with entities and topics extracted from news articles that are linked with Twitter messages (news-based enrichment). In previous work [12] we presented strategies for selecting appropriate news articles for enriching users’ Twitter activities.

A third dimension we investigate in the context of Twitter-based user modeling is given by *temporal constraints* that are considered when constructing the profiles (see Table 1). First, we study the nature of user profiles created within specific time periods. For example, we compare profiles constructed by exploiting the complete (long-term) user history with profiles that are based only on Twitter messages published within a certain week (short-term). Second, we examine

certain time frames for creating the profiles. For example, we explore the differences between user profiles created on the weekends with those created during the week to detect temporal patterns that might help to improve personalization within certain time frames.

By selecting and combining the different design dimensions and alternatives we obtain a variety of different user modeling strategies that will be analyzed and evaluated in this paper.

3.1 Twitter-based User Modeling Framework

We implemented the profiling strategies as a Twitter-based user modeling framework that is available via the supporting website of this paper [13]. Our framework features three main components that are essential for realizing the above user modeling strategies.

1. **Semantic Enrichment.** Given the content of Twitter messages we extract entities and topics to better understand the semantics of Twitter activities. Therefore we utilize OpenCalais², which allows for the detection and identification of 39 different types of entities such as persons, events, products or music groups and moreover provides unique URIs for identified entities as well as for the topics so that the meaning of such concepts is well defined.
2. **Linkage.** To further enrich the semantics of Twitter messages we implemented several strategies that link tweets with external Web resources. Previous research states that more than 85% of the Twitter activities are related to news events [3]. We thus developed linking strategies that detect whether and, if so, to which news article a tweet refers to – even for tweets that do not explicitly provide a hyperlink to the corresponding news article. In [12] we evaluated these strategies and showed that they achieve 70-80% accuracy. Given the links between tweets and news articles, entities and topics extracted from articles can be propagated to the corresponding tweets to further contextualize and enhance the semantics of Twitter activities.
3. **User Modeling.** Based on the semantic enrichment and the linkage with external news articles, our framework provides methods for generating hashtag-based, entity-based, and topic-based profiles that might adhere to specific temporal constraints (see above).

4 Analysis of Twitter-based User Profiles

To understand how the different user modeling design choices influence the characteristics of the generated user profiles, we applied our framework to conduct an in-depth analysis on a large Twitter dataset. The main research questions to be answered in this analysis can be summarized as follows.

1. How do the different user modeling strategies impact the *characteristics* of Twitter-based user profiles?
2. Which *temporal characteristics* do Twitter-based user profiles feature?

² <http://www.opencalais.com>

4.1 Data Collection and Data Set Characteristics

Over a period of more than two months we crawled Twitter information streams of more than 20,000 users. Together, these people published more than 10 million tweets. To allow for linkage of tweets with news articles we also monitored more than 60 RSS feeds of prominent news media such as BBC, CNN or New York Times and aggregated the content of 77,544 news articles. The number of Twitter messages posted per user follows a power-law distribution. The majority of users published less than 100 messages during our observation period while only a small fraction of users wrote more than 10,000 Twitter messages and one user produced even slightly more than 20,000 tweets (no spam). As we were interested in analyzing also temporal characteristics of the user profiles, we created a sample of 1619 users, who contributed at least 20 tweets in total and at least one tweet in each month of our observation period. This sample dataset contained 2,316,204 tweets in total.

We processed each Twitter message and each news article via the semantic enrichment component of our user modeling framework to identify topics and entities mentioned in the tweets and articles (see Section 3.1). Further, we applied two different linking strategies and connected 458,566 Twitter messages with news articles of which 98,189 relations were explicitly given in the tweets by URLs that pointed to the corresponding news article. The remaining 360,377 relations were obtained by comparing the entities that were mentioned in both news articles and tweets as well as by comparing the timestamps. In previous work we showed that this method correlates news and tweets with an accuracy of more than 70% [12]. Our hypothesis is that – regardless whether this enrichment method might introduce a certain degree of noise – it impacts the quality of user modeling and personalization positively.

4.2 Structural Analysis of Twitter-based Profiles

To validate our hypothesis and explore how the exploitation of linked external sources influences the characteristics of the profiles generated by the different user modeling strategies, we analyzed the corresponding profiles of the 1619 users from our sample. In Figure 1 we plot the number of distinct (types of) concepts in the topic- and entity-based profiles and show how this number is influenced by the additional news-based enrichment.

For both types of profiles the enrichment with entities and topics obtained from linked news articles results in a higher number of distinct concepts per profile (see Fig. 1(a) and 1(b)). Topic-based profiles abstract much stronger from the concrete Twitter activities than entity-based profiles. In our analysis we utilized the OpenCalais taxonomy consisting of 18 topics such as politics, entertainment or culture. The tweet-only-based user modeling strategy, which exploits merely the semantics attached to tweets, fails to create profiles for nearly 100 users (6.2%, topic-based) as for these users none of the tweets can be categorized into a topic. By enriching the tweets with topics inferred from the linked news articles we better understand the semantics of Twitter messages and succeed in creating more valuable topic-based profiles for 99.4% of the users.

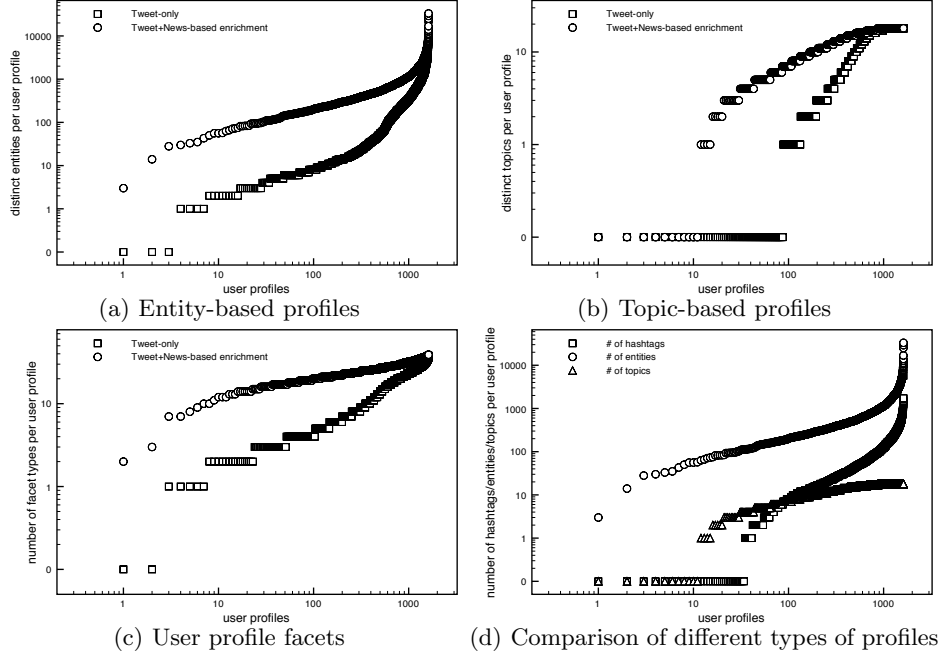


Fig. 1. Comparison between different user modeling strategies with tweet-only-based or news-based enrichment.

Further, the number of profile facets, i.e. the type of entities (e.g. person, location or event) that occur in the entity-based profiles, increases with the news-based semantic enrichment (see Fig. 1(c)). While more than 400 twitter-based profiles (more than 25%) feature less than 10 profile facets and often miss entities such as movies or products a user is concerned with, the news-based enrichment detects a greater variety of entity types. For more than 99% of the entity-based profiles enriched via news articles, the number of distinct profile facets is higher than 10.

A comparison of the entity- and topic-based user modeling strategies with the hashtag-based strategy (see Fig. 1(d)) shows that the variety of entity-based profiles is much higher than the one of hashtag-based profiles. While the entity-based strategy succeeds to create profiles for all users in our dataset, the hashtag-based approach fails for approximately 90 users (5.5%) as the corresponding people neither made use of hashtags nor re-tweeted messages that contain hashtags. Entity-based as well as topic-based profiles moreover make the semantics more explicit than hashtag-based profiles. Each entity and topic has a URI which defines the meaning of the entity and topic respectively.

The advantages of well-defined semantics as exposed by the topic- and entity-based profiles also depend on the application context, in which these profiles are used. The results of the quantitative analysis depicted in Fig. 1 show that entity- and topic-based strategies allow for higher coverage regarding the number of users, for whom profiles can be generated, than the hashtag-based strategy.

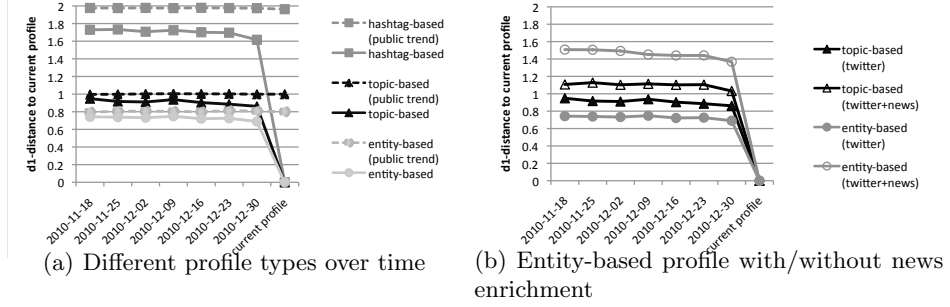


Fig. 2. Temporal evolution of user profiles: average d_1 -distance of current individual user profiles with corresponding profiles in the past.

Further, semantic enrichment by exploiting news articles (implicitly) linked with tweets increases the number of entities and topics available in the profiles significantly and improves the variety of the profiles (the number of profile facets).

4.3 Temporal Analysis of Twitter-based Profiles

In the temporal analysis we investigate (1) how the different types of user profiles evolve over time and (2) which temporal patterns occur in the profiles. Regarding temporal patterns we, for example, examine whether profiles generated on the weekends differ from those generated during the week. Similar to the click-behavior analysis by Liu et al. [14], we apply the so-called d_1 -distance for measuring the difference between profiles in vector representation.

$$d_1(\mathbf{p}_x(u), \mathbf{p}_y(u)) = \sum_i |p_{x,i} - p_{y,i}| \quad (1)$$

The higher $d_1(\mathbf{p}_x(u), \mathbf{p}_y(u)) \in [0..2]$ the higher the difference of the two profiles $\mathbf{p}_x(u)$ and $\mathbf{p}_y(u)$ and if two profiles are the same then $d_1(\mathbf{p}_x(u), \mathbf{p}_y(u)) = 0$. Figure 2 depicts the evolution of profiles over time. It shows the average d_1 -distance of the current user profiles with the profiles of the same users created based on Twitter activities performed in a certain week in the past. As suggested in [14], we also plotted the distance of the current user-specific profile with the *public trend* (see Fig. 2(a)), i.e. the average profile of the corresponding weeks.

For the three different profile types we observe that the d_1 -distance slightly decreases over time. For example, the difference of current profiles (first week of January 2011) with the corresponding profiles generated at the beginning of our observation period (in the week around 18th November 2010) is the highest while the distance of current profiles with profiles computed one week before (30th December 2010) is the lowest. It is interesting to see that the distance of the current profiles with the public trend (i) is present for all types of profiles and (ii) is rather constant over time. This suggests (i) a certain degree of individualism in Twitter and (ii) reveals that the people in our sample follow different trends rather than being influenced by the same trends.

Hashtag-based profiles exhibit the strongest changes over time as the average d_1 -distance to the current profile is constantly higher than for the topic- and entity-based profiles. Figure 2(b) discloses that entity-based profiles change

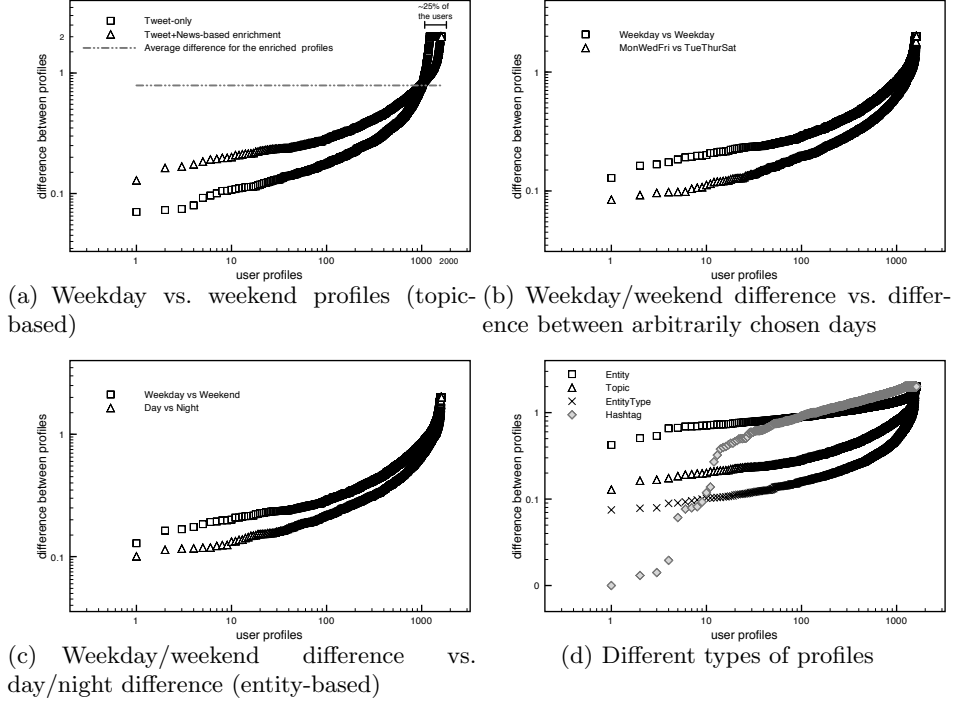


Fig. 3. Temporal patterns: comparison between weekend and weekday profiles by means of d_1 -distance.

stronger over time than topic-based profiles when news-based enrichment is enabled. When merely analyzing Twitter messages one would come to a different (possibly wrong) conclusion (see Fig. 2(a)).

Figure 3 illustrates temporal patterns we detected when analyzing the individual user profiles. In particular, we investigate how profiles created on the weekends differ from profiles (of the same user) created during the week. For topic-based profiles generated solely based on Twitter messages, it seems that for some users the weekend and weekday profiles differ just slightly while for 24.9% of the users the d_1 -distance of the weekend and weekday profile is maximal (2 is the maximum possible value, see Fig. 3(a)). The news-based enrichment reveals however that the difference of weekend and weekday profiles is a rather common phenomenon: the curve draws nearer to the average difference (see dotted line); there are less extrema, i.e. users for whom the d_1 -difference is either very low or very high. Hence, it rather seems that the tweets alone are not sufficient to get a clear understanding of the users concerns and interests.

Fig. 3(b) further supports the hypothesis that weekend profiles differ significantly from weekday profiles. The corresponding distances $d_1(\mathbf{p}_{\text{weekend}}(u), \mathbf{p}_{\text{weekday}}(u))$ are consistently higher than the differences of profiles generated on arbitrarily chosen days during the week. This *weekend pattern* is more significant than differences between topic-based profiles generated based on Twitter messages that are either posted during the evening (6pm-3am) or during the day

(9am-5pm) as shown in Fig. 3(c). Hence, the individual topic drift – i.e. change of topics individual users are concerned with – between day and evening/night seems to be smaller than between weekdays and weekends.

The weekend pattern is coherent over the different types of profiles. Different profile types however imply different drift of interests or concerns between weekend and weekdays (see Fig. 3(d)). Hashtag-based and entity-based profiles change most while the types of entities people refer to (persons, products, etc.) do not differ that strongly. When zooming into the individual entity-based profiles we see that entities related to leisure time and entertainment become more important on the weekends.

The temporal analysis thus revealed two important observations. First, user profiles change over time: the older a profile the more it differs from the current profile of the user. The actual profile distance varies between the different types of profiles. Second, weekend profiles differ significantly from weekday profiles.

5 Exploitation of User Profiles for Personalized News Recommendations

In this section we further analyze the different user modeling strategies, measure their impact on recommending news and try to exploit the temporal patterns to improve recommendation quality. We thus answer the following questions.

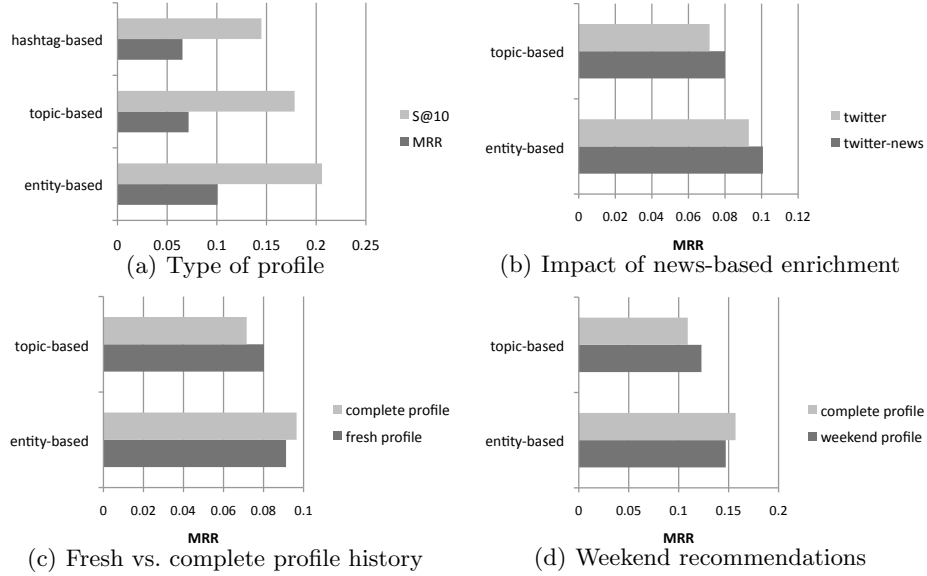
1. To which degree are the profiles created by the different user modeling strategies appropriate for recommending news?
2. Can the identified (temporal) patterns be applied to improve recommendation accuracy?

5.1 News Recommender System and Evaluation Methodology

Recommending news articles is a non-trivial task as the news items, which are going to be recommended, are *new* by its very nature, which makes it difficult to apply collaborative filtering methods, but rather calls for content-based or hybrid approaches [14]. Our main goal is to analyze and compare the applicability of the different user modeling strategies in the context of news recommendations. We do not aim to optimize recommendation quality, but are interested in comparing the quality achieved by the same recommendation algorithm when inputting different types of user profiles. Therefore we apply a lightweight content-based algorithm that recommends items according to their cosine similarity with a given user profile. We thus cast the recommendation problem into a search and ranking problem where the given user profile, which is constructed by a specific user modeling strategy, is interpreted as query.

Definition 2 (Recommendation Algorithm). *Given a user profile $\mathbf{p}(u)$ in vector representation and a set of candidate news items $N = \{\mathbf{p}(n_1), \dots, \mathbf{p}(n_n)\}$, which are represented via profiles using the same vector representation, the recommendation algorithm ranks the candidate items according to their cosine similarity to $\mathbf{p}(u)$:*

$$sim_{cosine}(\mathbf{p}(u), \mathbf{p}(n_i)) = \frac{\mathbf{p}(u) \cdot \mathbf{p}(n_i)}{\|\mathbf{p}(u)\| \cdot \|\mathbf{p}(n_i)\|} \quad (2)$$

**Fig. 4.** Results of news recommendation experiment.

Given the Twitter and news media dataset described in Section 4.1, we considered the last week of our observation period as the time frame for computing recommendations. The ground truth of news articles, which we consider as *relevant* for a specific user u , is obtained via the Twitter messages (including retweets) posted by u in this week that explicitly link to a news article published by BBC, CNN or New York Times. We thereby identified, on average, 5.5 relevant news articles for each of the 1619 users from our sample. For less than 10% of the users we found more than 20 relevant articles. The candidate set of news articles, which were published within the recommendation time frame, contained 5529 items. We then applied the different user modeling strategies together with the above algorithm (see Def. 2) and set of candidate items to compute news recommendations for each user. The user modeling strategies were only allowed to exploit tweets published before the recommendation period. The quality of the recommendations was measured by means of *MRR* (Mean Reciprocal Rank), which indicates at which rank the first item relevant to the user occurs on average, and *S@k* (Success at rank k), which stands for the mean probability that a relevant item occurs within the top k of the ranking. We tested statistical significance of our results with a two-tailed *t*-Test where the significance level was set to $\alpha = 0.01$ unless otherwise noted.

5.2 Results

The results of the news recommendation experiment are summarized in Fig. 4 and validate findings of our analysis presented in Section 4. Entity-based user modeling (with news-based enrichment), which produces according to the quantitative analysis (see Fig. 1) the most valuable profiles, allowed for the best recommendation quality and performed significantly better than hashtag-based user modeling (see Fig. 4(a)). Topic-based user modeling also performed better

than the hashtag-based strategy – regarding S@10 the performance difference is significant. Since the topic-based strategy models user interests within a space of 18 different topics (e.g., politics or sports), it further required much less run-time and memory for computing user profiles and recommendations than the hashtag- and entity-based strategies, for which we limited dimensions to the 10,000 most prominent hashtags and entities respectively.

Further enrichment of topic- and entity-based profiles with topics and entities extracted from linked news articles, which results in profiles that feature more facets and information about users’ concerns (cf. Section 4.2), also results in a higher recommendation quality (see Fig. 4(b)). Exploiting both tweets and linked news articles for creating user profiles improves MRR significantly ($\alpha = 0.05$). In Section 4.3 we observed that user profiles change over time and that recent profile information approximates future profiles slightly better than old profile information. We thus compared strategies that exploited just recent Twitter activities (two weeks before the recommendation period) with the strategies that exploit the entire user history (see Fig. 4(c)). For the topic-based strategy we see that *fresh* user profiles are more applicable for recommending news articles than profiles that were built based on the entire user history. However, entity-based user modeling enables better recommendation quality when the complete user history is applied. Results of additional experiments [13] suggest that this is due to the number of distinct entities that occur in entity-based profiles (cf. Fig. 1): long-term profiles seem to refine preferences regarding entities (e.g. persons or events) better than short-term profiles.

In Section 4.3 we further observed the so-called *weekend pattern*, i.e. user profiles created based on Twitter messages published on the weekends significantly differ from profiles created during the week. To examine the impact of this pattern on the accuracy of the recommendations we focused on recommending news articles during the weekend and compared the performance of user profiles created just by exploiting weekend activities with profiles created based on the complete set of Twitter activities (see Fig. 4(d)). Similarly to Fig. 4(c) we see again that the entity-based strategy performs better when exploiting the entire user history while the topic-based strategy benefits from considering the weekend pattern. For the topic-based strategy recommendation quality with respect to MRR improves significantly when profiles from the weekend are applied to make recommendations during the weekend.

6 Conclusions

In this paper we developed a user modeling framework for Twitter and investigated how the different design alternatives influence the characteristics of the generated user profiles. Given a large dataset consisting of more than 2 million tweets we created user profiles and revealed several advantages of semantic entity- and topic-based user modeling strategies, which exploit the full functionality of our framework, over hashtag-based user modeling. We saw that further enrichment with semantics extracted from news articles, which we correlated with the users’ Twitter activities, enhanced the variety of the constructed profiles and improved accuracy of news article recommendations significantly.

Further, we analyzed the temporal dynamics of the different types of profiles. We observed how profiles change over time and discovered temporal patterns such as characteristic differences between weekend and weekday profiles. We also showed that the consideration of such temporal characteristics is beneficial to recommending news articles when dealing with topic-based profiles while for entity-based profiles we achieve better performance when incorporating the entire user history.

In future work we will further research the temporal specifics of entity-based profiles. First results [13] suggest that users refer to certain types of entities (e.g., persons) more consistently over time than to others (e.g., movies or events). Therefore we plan to explore whether knowledge regarding the temporal validity of the different profile facets can further leverage personalization quality.

References

1. Lerman, K., Ghosh, R.: Information contagion: an empirical study of spread of news on Digg and Twitter social networks. In: Proc. of 4th Int. Conf. on Weblogs and Social Media (ICWSM), AAAI Press (2010)
2. Sakaki, T., Okazaki, M., Matsuo, Y.: Earthquake shakes Twitter users: real-time event detection by social sensors. In: Proc. of 19th Int. Conf. on World Wide Web (WWW), ACM (2010) 851–860
3. Kwak, H., Lee, C., Park, H., Moon, S.: What is Twitter, a social network or a news media? In: Proc. of 19th Int. Conf. on World Wide Web (WWW), ACM (2010) 591–600
4. Weng, J., Lim, E.P., Jiang, J., He, Q.: TwitterRank: Finding topic-sensitive influential Twitterers. In Davison et al., eds.: Proc. of 3rd Int. Conf. on Web Search and Web Data Mining (WSDM), ACM (2010) 261–270
5. Cha, M., Haddadi, H., Benevenuto, F., Gummadi, P.K.: Measuring User Influence in Twitter: The Million Follower Fallacy. In Cohen, W.W., Gosling, S., eds.: Proc. of 4th Int. Conf. on Weblogs and Social Media (ICWSM), AAAI Press (2010)
6. Dong, A., Zhang, R., Kolar, P., Bai, J., Diaz, F., Chang, Y., Zheng, Z., Zha, H.: Time is of the essence: improving recency ranking using twitter data. In: Proc. of 19th Int. Conf. on World Wide Web (WWW), ACM (2010) 331–340
7. Chen, J., Nairn, R., Nelson, L., Bernstein, M., Chi, E.: Short and tweet: experiments on recommending content from information streams. In: Proc. of 28th Int. Conf. on Human factors in Computing Systems (CHI), ACM (2010) 1185–1194
8. Laniado, D., Mika, P.: Making sense of Twitter. In Patel-Schneider et al., eds.: Proc. of 9th Int. Semantic Web Conference (ISWC), Shanghai, China. Volume 6496 of LNCS. Springer (2010) 470–485
9. Huang, J., Thornton, K.M., Efthimiadis, E.N.: Conversational Tagging in Twitter. In Chignell, M.H., Toms, E., eds.: Proc. of 21st Conf. on Hypertext and Hypermedia (HT), ACM (2010) 173–178
10. Koren, Y.: Collaborative filtering with temporal dynamics. In Elder et al., eds.: Proc. of 15th Int. Conf. on Knowledge Discovery and Data Mining (KDD), Paris, France, ACM (2009) 447–456
11. Rowe, M., Stankovic, M., Laublet, P.: Mapping Tweets to Conference Talks: A Goldmine for Semantics. In Passant et al., eds.: Workshop on Social Data on the Web (SDoW), co-located with ISWC '10, Shanghai, China. Volume 664., CEUR-WS.org (2010)
12. Abel, F., Gao, Q., Houben, G.J., Tao, K.: Semantic Enrichment of Twitter Posts for User Profile Construction on the Social Web. Technical report, submitted to: Extended Semantic Web Conference (ESWC), Heraklion, Greece (2011)
13. Abel, F., Gao, Q., Houben, G.J., Tao, K.: Supporting website: code, datasets and additional findings. (2011) <http://wis.ewi.tudelft.nl/umap2011/>.
14. Liu, J., Dolan, P., Pedersen, E.R.: Personalized news recommendation based on click behavior. In Rich et al., eds.: Proc. of 14th Int. Conf. on Intelligent User Interfaces (IUI), ACM (2010) 31–40