TUMS: Twitter-based User Modeling Service

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

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

Abstract. Twitter is today's most popular micro-blogging service on the Social Web. As people discuss various fresh topics, Twitter messages (tweets) can tell much about the current interests and concerns of a user. In this paper, we introduce TUMS, a Twitter-based User Modeling Service, that infers semantic user profiles from the messages people post on Twitter. It features topic detection and entity extraction for tweets and allows for further enrichment by linking tweets to news articles that describe the context of the tweets. TUMS is made publicly available as a Web application. It allows end-users to overview Twitter-based profiles in a structured way and allows them to see in which topics or entities a user was interested in a specific point in time. Furthermore, it provides Twitter-based user profiles in RDF format and allows applications to incorporate these profiles in order to adapt their functionality to the current interests of a user. TUMS is available via: http://wis.ewi.tudelft.nl/tums/

Key words: user modeling, twitter, semantic enrichment, service

1 Introduction

Applications that aim for personalization and would like to adapt their functionality to the current interests and demands of a user require information about their users [1]. User-adaptive systems suffer from cold-start and sparsity problems [2]. For example, when systems have to deal with new users or fresh content, those systems require user profile information that allows for estimating the interests of the users. In this paper, we present a Twitter-based user modeling service (TUMS) that exploits users's Twitter activities to infer semantically meaningful profiles.

Since Twitter was launched in 2007, it became the most popular micro-blogging service, with 190 million users who post more than 65 million posts every day[1]. As people are not limited to a certain domain, they can discuss various topics of which most of the topics are related to news which could also be found in mainstream news articles [3]. There is also an active research community on Twitter studying topics including social network analysis [3, 4], community and user influence mining [5–7], recommendations of URLs [8] or sentiment analysis [9, 10]. Inferring user profiles from individual Twitter activities is a hard

[1] http://techcrunch.com/2010/06/08/twitter-190-million-users/

problem as tweets are limited to 140 characters which makes the deduction of semantics difficult. Rowe et al. [11] propose to exploit the context in which tweets have been published. In particular, they propose to map tweets to events and exploit semantic descriptions of these events to clarify the semantic meaning of tweets. In previous work, we proposed strategies to automatically map Twitter messages to related news articles which allows us to exploit the news articles to enrich the semantics of individual tweets [12]. This enrichment builds the basis for the user modeling strategies [13] that are made available via the TUMS service which we present in this paper.

The TUMS service allows end-users to inspect Twitter-based profiles and enables other applications to re-use these profiles. People can overview their personal Twitter activities or profiles of other users to explore the topics those users were concerned with in the past. Entity-based, topic-based and hashtag-based tag clouds allow to quickly grasp the content of the profiles and charts further visualize the profiles. Visualizing the profiles is important to make people aware of what can be inferred from their Twitter activities so that they can reconsider how they publish tweets on Twitter. TUMS is also of interest for other services on the Social Web as it enables them to consume Twitter-based profiles in RDF format. Profiles can be used for personalization and are particularly interesting for other applications that suffer from sparsity problems (e.g. services that cannot collect sufficient data about their users) and services that are interested in “realtime” or very fresh profile information.

In the following section we summarize related work before we introduce TUMS in detail in Section 3. We outline the architecture of our service, the user modeling strategies featured by TUMS and present the graphical user interface as well as the service API. In Section 4 we conclude and give an outlook on our plans for future work.

2 Related Work

With the advent of Semantic Web technologies and appropriate vocabularies such as FOAF [14], SIOC [15] or the Weighted Interest vocabulary [16], re-use of user profiles is becoming easier nowadays. Research on generic user modeling services [17], mediating user models [18], identifying users across system boundaries [19] and cross-system user modeling and personalization [20, 21] further supports the re-use of user profiles in different application contexts. In this paper, we introduce a service that generates user profiles by exploiting Twitter and allows for applying these profiles in other applications.

Research on Twitter often focuses on analyzing large fractions of the Twitter network to study information propagation patterns [3, 22, 23] or to identify influential users [6, 7]. Dong et al. [24] 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 [8].

Yet, there exists little research on analyzing the semantics of individual tweets and exploiting Twitter as a source for modeling user interests. Rowe et al. [11] proposed to exploit contextual information to enrich the semantics of Twitter messages. In previous work, we followed this suggestion and linked tweets to related Web resources [12]. Given this semantic enrichment, we proposed user modeling strategies that allow for recommending news articles. In this paper, we make those enrichment and user modeling strategies available to the public and present TUMS, a Twitter-based User Modeling Service.

3 TUMS: Twitter-based User Modeling Service

TUMS is a user modeling service that exploits the Twitter messages of individual users to infer and provide user interest profiles. TUMS targets two types of consumers: (i) end-users who would like to overview and inspect their profiles (or profiles of friends) in a structured way and (ii) applications that would like to incorporate the RDF-based user profiles in order to adapt their functionality to the current interests of a user. In this section, we first present the architecture of TUMS before we summarize the user modeling capabilities of TUMS (see Section 3.2

3.1 Architecture

TUMS aims to to deliver both a human and a machine readable representation of the user characteristics that can be deduced from the posts a user published on Twitter. The only input required by TUMS is the Twitter username of the person that should be profiled. Figure. 1 depicts the architecture of the TUMS service. TUMS is composed of three modules: (i) data crawling and storage, (ii) Twitter-based user modeling module, (iii) a Web service that enables applications to query for profiles and (iv) a graphical user interface that allows applications to consume profiles.

Crawler and back-end From a user's perspective the TUMS functionality is very easy to use as TUMS just requires a Twitter username as input. Given a username, TUMS initiates a pipeline for crawling, processing, storing and analyzing the Twitter data that is publicly available for the corresponding user. TUMS features a tweets and news crawler: the tweets crawler aggregates Twitter posts of the given user and stores the tweets in the TUMS data repository. The news crawler continuously monitors traditional news media. At the moment, TUMS monitors 62 RSS feeds published by mainstream news publishers such as CNN, New York Times, and BBC and aggregates the full content of each news article by making use of Boilerpipe[2] Once a

[2] `http://code.google.com/p/boilerpipe/`, a Java library that detects supplemental text (e.g. advertisements or menus) in a website and allows for the extraction of the main content.

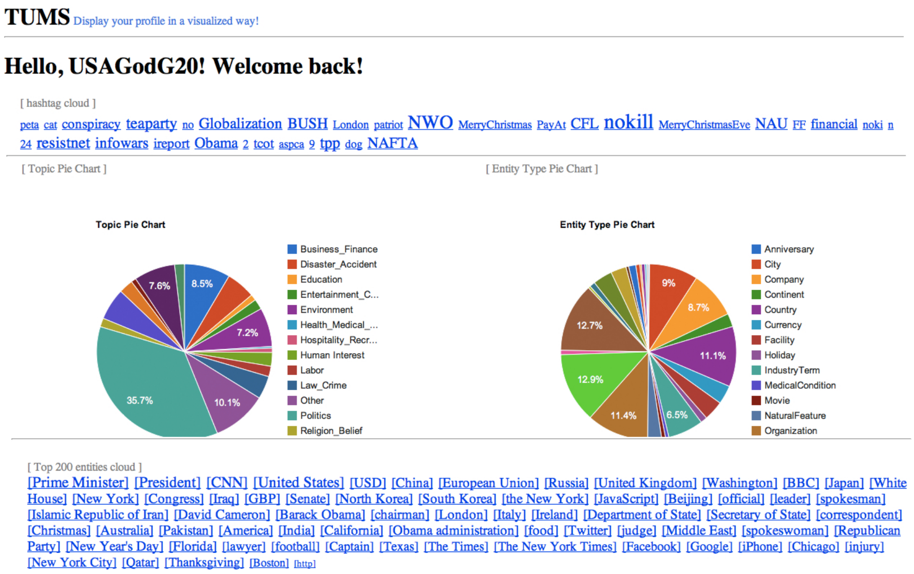

Fig. 1. The architecture of TUMS

new username is provided through the graphical interface, TUMS will continuously start monitoring the tweets of the corresponding user. Whenever a new tweet is observed, the tweet will get processed to enrich the semantics of the tweet (see below). Tweets, news articles and semantic metadata, which is extracted from tweets and news, are stored in the data repository which builds the basis for user modeling.

User modeling Given the raw data crawled from Twitter, TUMS aims to infer user interest profiles that adhere to the Friend-Of-A-Friend (FOAF) [14] vocabulary and the Weighted Interest vocabulary [16]. Hence, given the raw text of tweets, TUMS will output profiles which specify how much a user is interested into a certain topic. For example:

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix wi: <http://purl.org/ontology/wi/core#> .
@prefix wo: <http://purl.org/ontology/wo/core#> .
@prefix ex: <http://example.org/> .
@prefix dbpedia: <http://dbpedia.org/resource/>

<http://twitter.com/taubau>
  a foaf:Person ;
  foaf:name "Ke Tao";
  wi:preference [
     a wi:WeightedInterest ;
     wi:topic dbpedia:Jazz ;
     wo:weight [
        a wo:Weight ;
```

```
wo:weight_value 10.0 ;
wo:scale ex:AScale
] .
```

The above profile snippet expresses that a user (*http://twitter.com/taubau*), whose name is “Ke Tao” is interested into jazz (*dbpedia:Jazz*). This interest is further classified via a weight (*wo:weight_value*) on a certain scale (*wo:scale*). In order to infer interests of a user from the tweets this user posted on Twitter, TUMS performs four main steps: (1) it notifies the tweet crawler to start collecting tweets, (2) it enriches the semantics of the user's tweets by categorizing the topic of a tweet and extracting hashtags and other entities (e.g. persons, locations) from the tweet, (3) if possible, it links the tweet to related news articles to further enrich the semantics of a tweet with semantics extracted from the news article and (4) it applies different user modeling strategies (see Section 3.2) to generate user profile information.
The first three steps will be executed whenever a new Twitter message is observed while the last step, the actual user modeling, is executed on query time, i.e. when a user (via the graphical interface) or client application (via the Web service) requests profile information.

Graphical Interface and RDF Endpoint Given the data aggregated, processed and stored by the above modules, the Twitter-based user profiles can be retrieved by visiting the graphical user interface or by invoking the TUMS Web service API. The graphical interface enables end-users to overview their profiles by means of tag clouds (topic-based, hashtag-based and entity-based clouds) and diagrams (see Section 3.3). The REST-like Web service interface outputs user profiles in RDF using FOAF [14] and Weighted Interest vocabulary [16] as depicted in the above example.

3.2 User Modeling Strategies

The goal of TUMS and the user modeling strategies in particular is to generate user interest profiles that conform to the following model.

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.

To generate such profiles, we developed a set of user modeling strategies [13] that vary in four design dimensions: (i) the type of profiles created by the strategies, (ii) the weighting scheme, (iii) the data sources exploited to further enrich the Twitter-based profiles and (iv) temporal constraints that are considered when constructing the profiles (see Table 1). The generic model for profiles representing users is specified in Definition 1.

As listed in Table 1, TUMS allows for three *types* of profiles that differ with respect to the type of concepts $c \in C$ for which we specify an interest weight (see

design dimension	design alternatives
profile type	(i) hashtag-based, (ii) topic-based or (iii) entity-based
weighting schemes	(i) term frequency (TF) or (ii) $TF \times IDF$
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 (*weekend*, *night*, etc.) or (iii) no constraints

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

Definition 1): entity-based, topic-based and hashtag-based profiles. For entity-based profiles, we differentiate between 39 types of entities such as persons, organizations, cities, countries, events, or medical conditions. The topic-based profiles are rather broad and abstract from the concrete content as we map the tweets to 19 static topics such as sports, politics or music.

TUMS allows for different methods as weighting scheme $w(u, c)$. For example, using term frequency (TF), the weight of a concept is determined by the number of Twitter activities in which user u refers to concept c. In a hashtag-based profile $w(u, \#technology) = 5$ means that u published five Twitter messages that mention "#technology". Other weighting methods such as $TF \times IDF$ are possible as well and a detailed comparison of different weighting schemes is planed for future work. The resulting user profiles will be normalized 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 $\boldsymbol{p}(u)$ we refer to $P(u)$ in its vector space model representation, where the value of the i-th dimension refers to $w(u, c_i)$.

Twitter messages posted by a user u may refer to external resources. A user can explicitly link to other Web resources in her Twitter message or she could discuss topics and events that are, for example, discussed in mainstream news articles as well. TUMS aims to also take these external resources into account when constructing entity-based and topic-based user profiles (see semantic enrichment in Table 1). In particular, profiles are enriched with entities and topics extracted from news articles that are linked with Twitter messages. In previous work, we have done the study on selecting appropriate news articles for enriching tweets [12]. We revealed that for Twitter messages, which do not explicitly link to external Web resources we can find related news articles that report about the same event as the Twitter message with a high precision of more than 70%.

Temporal constraints are also considered as the fourth dimension of the profiles (see Table 1). By specifying temporal constraints, client applications can, for example, retrieve the latest profile of a user or a profile that is only based on Twitter activities which a user performed on weekends. In previous work [13], we revealed, for example, that there are significant differences between user profiles created on the weekends with those created during the week.

By selecting and combining the different design dimensions and alternatives, TUMS can generate different profiles in the form of visualized charts and RDF triples.

Fig. 2. Graphical user interface on the start page of TUMS

3.3 Graphical User Interface of TUMS

End-users can easily access TUMS via its graphical user interface which requires only a Twitter username as input and is thus easy to use. Figure 2 shows the start screen that is displayed when users are accessing TUMS. In the given example, the profile of a user named "USAGodG20" should be returned. By clicking on the "Get Tweet Profile" button the user will be directed to the next page that allows the user to overview and inspect the profile of USAGodG20 (see Fig. 3).

TUMS can represent user profiles by means of different visualizations, including hashtag-based clouds, pie charts visualizing topic-based profiles and entity types referenced by a user, and entity-based clouds (see Fig. 3). In the following, we examine the implemented visualization features of TUMS and describe them by example results that we obtain for requesting the profile of the user "USAGodG20".

Hashtag-based clouds The hashtag-based clouds represent the hashtag-based profile of a user and thus visualize the hashtags which the corresponding user mentioned in her Twitter messages. Different hashtags will be displayed in different font size depending on their weight within the hashtag-based profile (cf. $w(u, c)$ in Definition 1). The higher the weight $w(u, c)$ within the profile $P(u)$ the bigger the font size. Hashtags displayed within the tag cloud are hyperlinks and direct the user to the Twitter search which will display the lastest tweets that mention the hashtag.

For example, at the top of Fig. 3 we see the hashtag cloud of "USAGodG20". The visualization of the hashtag-based profile tells us that "USAGodG20" is particularly interested in the news about globalization, NAFTA[3], and political leaders in the United States, e.g. Bush, Obama.

Pie charts for topic-based profiles One of the reasons for the popularity of Twitter lies in the idea that anyone can contribute and discuss her/his opinion about any topic. TUS generates topic-based profiles and the TUMS user interface provides a pie chart that displays the relative importance of a topic

[3] NAFTA stands for North American Free Trade Agreement. It is an agreement signed by the governments of Canada, Mexico, and the United States, creating a trilateral trade bloc in North America.

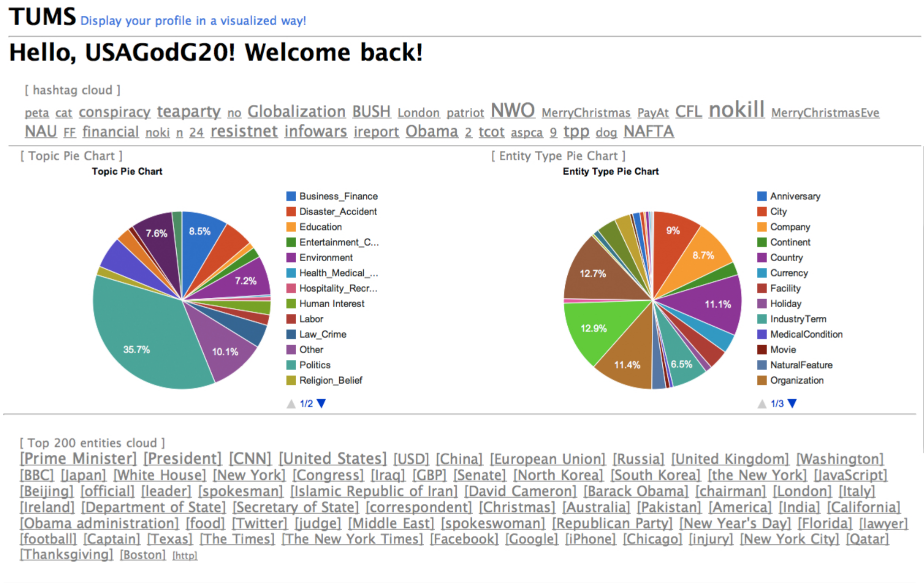

Fig. 3. Overview of user profile page of TUMS

within a user profile. Topic-based profiles feature 19 different dimensions that correspond to broad topics like politics or education. Given the pie chart visualization of a topic-based profiles, we can see about which topics a user has been published most.

Fig. 4 zooms into the pie chart of USAGodG20's profile and shows that 35.7% of his tweets are concerned with politics. This observation is consistent with the results we got from the hashtag-based cloud. Via the pie chart we can moreover compare this interest with interests for other topics. For example, we see that the user is also interested in "Business & Finance" (8.5%) but rather not interested in education or entertainment and culture.

Pie charts of referenced entity types During the user modeling process, entities are extracted from tweets and external sources for enriching the semantic of Twitter messages. TUMS extracts 39 types of the entities. The pie chart of entity types shows the importance of the entity types for a given user and can be inferred from the entity-based profiles.

The example pie chart generated based on the entity-based profile of USAGodG20 shows that USAGodG20 is often referring to persons (12.9%), organizations (11.4%), countries (11.1%) and cities (9%). Natural features (e.g. beach, ocean) or movies are less frequently mentioned.

Entity-based clouds With entity-based clouds, TUMS visualizes the entity-based profile of a user and thus lists the top entities mentioned by a user. It refines the entity type based pie chart overview and shows particular persons, locations, etc. in which a user is interested. As for the hashtag-based cloud,

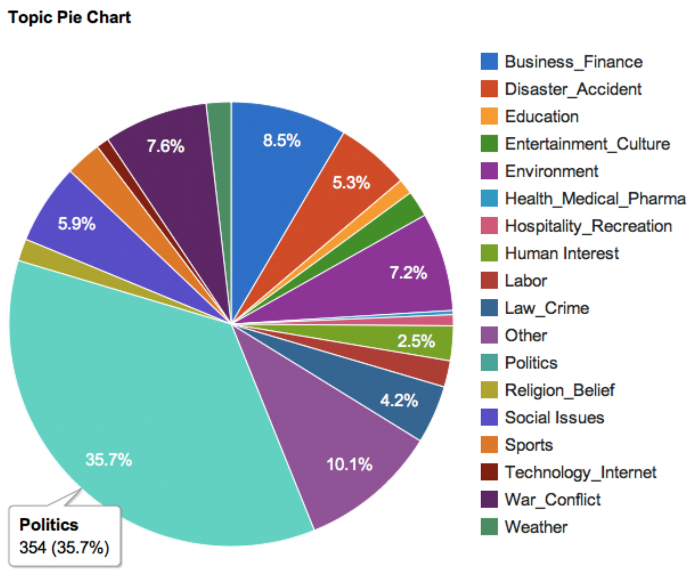

Fig. 4. The screenshot of the pie chart of topics

entities are represented in different font size according to their weight $w(u, c)$ in the entity-based profile (cf. Definition 1).
At the bottom of Fig. 3 we see the most important entities within the entity-based profile of USAGodG20. Entities with larger font size in the cloud, such as "Prime Minister", "President", "United States", "China", and "European Union" are words that are quite common within the political domain which confirms again the findings from the above profile visualizations.

3.4 RDF Endpoint of TUMS

User profiles are furthermore made available via Web service in RDF format. Via HTTP content negotiation[4] or by specifying the format in the HTTP query string (*format=rdf*), client applications are enabled to request the different types of profiles generated by TUMS in RDF. The URL pattern that represents a query for profile information is defined as follows:

```
.../profile/[username]/[profile type]/[weighting]/[enrichment]/?[temporal
constraints]
```

Hence, in addition to the *username*, which refers to the Twitter username of the person whose profile should be returned, there are four types of parameters which can be customized in the pattern. These four parameters correspond to the four dimensions in Table 1. The format and possible values of these parameters are explained below.

[4] http://tools.ietf.org/html/rfc2616#section-12

`profile type` The profile type parameter enables applications to specify which type of profile they would like to retrieve: hashtag-based, topic-based or entity-based. Given a user profile $P(u) = \{(c_1, w(u, c_1)), (c_2, w(u, c_2)), ...\}$ where $c_i \in C$, the profile type thus specifies the type of the concepts c_i. While topics and entities are represented via a URI by nature and therefore explicitly describe the semantic meaning of a topic and entity respectively, we try to clarify the semantic meaning of hashtags by referring to *#tagdef*[5]. Given the three types of profiles, the corresponding parameter values that can be used in the HTTP request are the following: *hashtag*, *topic*, and *entity*. For example, the hashtag-based profile can be retrieved as follows:

```
.../profile/USAGodG20/hashtag/
```

The RDF-formatted response to this HTTP request might be the following:

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix wi: <http://purl.org/ontology/wi/core#> .
@prefix wo: <http://purl.org/ontology/wo/core#> .
@prefix tums: <http://wis.ewi.tudelft.nl/tums/rdf/> .

<http://twitter.com/USAGodG20>
  a foaf:Person;
  foaf:name "John of God";
  wi:preference [
     a wi:WeightedInterest ;
     wi:topic <http://tagdef.com/conspiracy> ;
     wo:weight [
        a wo:Weight ;
        wo:weight_value 0.0757 ;
        wo:scale tums:Scale
        ] ;
  wi:preference [
     a wi:WeightedInterest ;
     wi:topic <http://tagdef.com/Globalization> ;
     wo:weight [
        a wo:Weight ;
        wo:weight_value 0.1576 ;
        wo:scale tums:Scale
        ] .
```

Hence, the hashtag-based profile of USAGodG20, which is in the graphical user interface represented by means of a tag cloud (see above), is represented using FOAF [14] and the Weighted Interest vocabulary [16]. For each concept (*wi:topic*) in the profile, the weight is clearly specified using *wo:weight_value*. TUMS normalizes the weights in a profile vector so that the sum of all weights in the profile is equal to 1. Correspondingly, the scale (*tums:Scale*) on which the weights are specified ranges from 0 to 1 (*wo:min_weight = 0*, *wo:max_weight = 1*, cf. Weighting Ontology[6]).

[5] http://tagdef.com
[6] http://purl.org/ontology/wo/core

`weighting` TUMS supports different weighting schemes for computing the weights in a user profile (cf. $w(u, c)$ in Definition 1). At the moment, TUMS supports (i) term frequency (TF) and (ii) TF multiplied by inverse document frequency ($TF \times IDF$). The corresponding parameters are *tf* and *tfidf* so that a request may have the following format:

```
.../profile/USAGodG20/hashtag/tf/
```

`enrichment` Currently, the possible options for this parameter are: (i) *tweet* or (ii) *tweetnews*. By default, TUMS extracts semantics from Twitter messages (*tweet*). However, developers can specify whether external sources should also be exploited to further enrich the semantics of tweets when constructing the profile (*tweetnews*). If no parameter is specified then TUMS aims to provide the best possible profile and thus will returned the profile that is further enriched with semantics extracted from related news articles (*tweetnews*). In future versions of TUMS we plan further options for adjusting the semantic enrichment such as means for selecting the target ontology of concepts that are referenced as *wi:topic* (e.g. developers should be enabled to decide whether these concepts should be mapped to DBpedia concepts). Furthermore, we plan to parametrize the linkage strategies that are used for the news-based enrichment (*tweetnews*).

`temporal contraints` Temporal constraints of profiles can be specified in the URL-based query as well. For specifying a certain time period, developers can use `start` and `end` to specify the start and end of the period that should be considered when generating the profiles. For example, the profile from December 1st 2010 to February 28th 2011 can be retrieved via:

```
.../profile/USAGodG20/hashtag/?start=2010-12-01&end=2011-02-28
```

Moreover, it is possible to specify pattern-like time constraints like weekend or night. They can be specified by adding further parameters to the query string. At the moment, TUMS allows for a `wk` parameter which can be *weekend* (Twitter activities on Saturdays or Sundays) or *weekday* (Monday till Friday) and a `dk` parameter which can be *night* (from 6pm till 6am) or *day* (from 6am till 6pm). For example:

```
.../hashtag/?start=2010-12-01&end=2011-02-28&wk=weekend&dn=night
.../hashtag/?wk=weekday
```

If no temporal constraints are specified then TUMS will create the profiles based on all available Twitter activities performed by the user.

3.5 Preliminary Evaluation of TUMS

Given tweets of 1619 users for which TUMS created user profiles, we randomly selected 50 entity and 50 topic assignments that were inferred by TUMS as part of the semantic enrichment. Ratings were done on a three-dimensional scale. For entity extraction, 3 (*perfect match*) means extracted entity has the name and

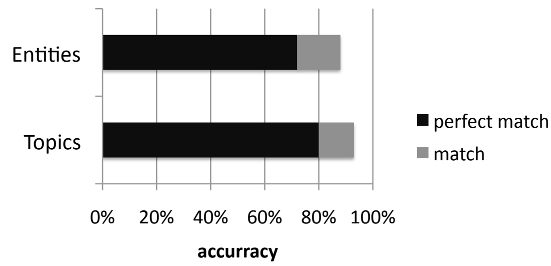

Fig. 5. Accuracy of entity extraction and topic detection.

type which exactly match with its meaning in the tweet; 2 (*match*) stands for that there might be some noisy words in the name of the entity, but the part other than the noise still matches with its real meaning, and the type the entity is correct; 1 (*not match*) denotes that the entity doesn't match the meaning in the tweet. For topic detection, the rating of 3 means the detected topic perfectly fits with the content of the tweet; the tweet rated as 2 has the content partly related to with the assigned topic; the wrong topic detection will be rated as 1.

Figure 5 lists the results and shows that, nearly 90% of the entity assignments and 70% percent topic detection are rated as 3 (*perfect match*) or 2 (*match*). Among these meaningful semantic enrichment, 82% of the entity assignments and 80% of the topic detection exactly match the meaning in the tweets.

3.6 Application Scenarios

There exist several applications that will benefit from using TUMS. In particular, systems that aims for personalization and are interested in fresh profile information such as news recommender systems might be interested in the Twitter-based profiles delivered by TUMS. Moreover, the visualization of the profiles can help adaptive systems to explain their users why certain adaptation decisions have been done, e.g. why an item was recommended.

Recommending Fresh Content Liu et al.[25] analyzed content-based recommender for Google News. They showed that interests in news topics such as technologies, politics, etc. change over time. Hence, the latest interests of a user are beneficial for building a personalization system. Given that most tweets on Twitter are related to news [3], Twitter seems to be an interesting source of information and TUMS the corresponding service that can support systems that are interested in current interests of a user. In [13] we report on first results in recommending news articles based on TUMS profiles. The TUMS profile can also be used for generating personalized email, recommending webpages, etc.

Explaining Recommendations The semantic user profiles and the visualizations of the profiles provided by TUMS can be applied to explain why recommendations have been generated. Such explanations may foster acceptance of recommendations. For example, for some e-shopping website like Amazon[7] it is common to give an explanation why a product was recommended. Using TUMS, recommender systems could link to the profile of a user that they used in order to compute the recommendation.

Authoring Support Services like Zemanta[8] or SMOB [26] enable authors of (micro)blogs to attach semantics to their posts. Since TUMS user profiles such as entity-based profiles can be used to support the user when annotating their blog and microblog posts with semantically meaningful URIs.

Privacy Awareness Twitter users may not know what can be derived from their Twitter activities. Hence, the graphical components of TUMS that allow for exploring the Twitter-based profiles in a more structured way can foster user awareness.

4 Conclusion and Future Work

In this paper, we present TUMS, a Twitter-based User Modeling Service, that generates semantic user profiles by exploiting Twitter messages. Based on functionality for enriching the semantics of tweets, TUMS features a variety of user modeling strategies that produce entity-based, topic-based or hashtag-based profiles. TUMS makes these profiles available to other applications via an RDF endpoint and allows end-users to explore profiles visually. Given a Twitter username, TUMS aggregates tweets from Twitter and starts monitoring the Twitter user. All tweets published by the user will be processed by the semantic enrichment module of TUMS which, for example, extracts entities from tweets and links tweets to related news articles to further enrich the semantics of a tweet. Given the semantically enriched tweets, TUMS can be queried to return semantic profiles for specific periods in time.

In our future work, we plan to support platforms that aim for personalization with TUMS user modeling capabilities. First experiments show that entity-based user profiles generated by TUMS, allow for high precision when recommending news articles [13]. Our ambition is to make this news recommender available as Web application using TUMS for generating user profiles.

Acknowledgements. This work is partially sponsored by the EU FP7 projects ImREAL (http://imreal-project.eu) and GRAPPLE (http://grapple-project.org).

References

1. Jameson, A.: Adaptive interfaces and agents. The HCI handbook: fundamentals, evolving technologies and emerging applications (2003) 305–330

[7] http://www.amazon.com
[8] http://www.zemanta.com

2. Brusilovsky, P., Kobsa, A., Nejdl, W., eds.: The Adaptive Web, Methods and Strategies of Web Personalization. In Brusilovsky, P., Kobsa, A., Nejdl, W., eds.: The Adaptive Web. Volume 4321 of Lecture Notes in Computer Science., Springer (2007)
3. Kwak, H., Lee, C., Park, H., Moon, S.: What is twitter, a social network or a news media? In: Proceedings of the 19th International Conference on World Wide Web (WWW '10), New York, NY, USA, ACM (2010) 591–600
4. Krishnamurthy, B., Gill, P., Arlitt, M.: A few chirps about twitter. In: Proceedings of the first workshop on Online social networks. WOSP '08, New York, NY, USA, ACM (2008) 19–24
5. Java, A., Song, X., Finin, T., Tseng, B.: Why we twitter: understanding microblogging usage and communities. In: Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis. WebKDD/SNA-KDD '07, New York, NY, USA, ACM (2007) 56–65
6. Weng, J., Lim, E.P., Jiang, J., He, Q.: Twitterrank: finding topic-sensitive influential twitterers. In Davison, B.D., Suel, T., Craswell, N., Liu, B., eds.: Proceedings of the Third International Conference on Web Search and Web Data Mining (WSDM '10), New York, NY, USA, ACM (2010) 261–270
7. 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.: Proceedings of the Fourth International Conference on Weblogs and Social Media (ICWSM '10), Washington, DC, USA, The AAAI Press (2010)
8. Chen, J., Nairn, R., Nelson, L., Bernstein, M., Chi, E.: Short and tweet: experiments on recommending content from information streams. In: Proceedings of the 28th international conference on Human factors in computing systems (CHI '10), New York, NY, USA, ACM (2010) 1185–1194
9. Gaffney, D.: #iranElection: quantifying online activism. In: Proceedings of the WebSci10: Extending the Frontiers of Society On-Line. (2010)
10. Diakopoulos, N.A., Shamma, D.A.: Characterizing debate performance via aggregated twitter sentiment. In: Proceedings of the 28th international conference on Human factors in computing systems. CHI '10, New York, NY, USA, ACM (2010) 1195–1198
11. Matthew Rowe and Milan Stankovic and Philippe Laublet: Mapping Tweets to Conference Talks: A Goldmine for Semantics. In Passant, A., , Breslin, J., Fernandez, S., Bojars, U., eds.: Workshop on Social Data on the Web (SDoW '10), colocated 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. In: Extended Semantic Web Conference (ESWC '11), Heraklion, Greece, Springer (May 2011)
13. Abel, F., Gao, Q., Houben, G.J., Tao, K.: Analyzing User Modeling on Twitter for Personalized News Recommendations. In: International Conference on User Modeling, Adaptation and Personalization (UMAP), Girona, Spain, Springer (July 2011)
14. Brickley, D., Miller, L.: FOAF Vocabulary Specification 0.91. Namespace document, FOAF Project, http://xmlns.com/foaf/0.1/ (November 2007)
15. Bojars, U., Breslin, J.G.: SIOC Core Ontology Specification. Namespace document, DERI, NUI Galway, http://rdfs.org/sioc/spec/ (January 2009)
16. Brickley, D., Miller, L., Inkster, T., Zeng, Y., Wang, Y., Damljanovic, D., Huang, Z., Kinsella, S., Breslin, J., Ferris, B.: The Weighted Interests Vocabulary 0.5. Namespace document, Sourceforge (September 2010)

17. Kobsa, A.: Generic user modeling systems. User Modeling and User-Adapted Interaction **11**(1-2) (2001) 49–63
18. Berkovsky, S., Kuflik, T., Ricci, F.: Mediation of user models for enhanced personalization in recommender systems. User Modeling and User-Adapted Interaction (UMUAI) **18**(3) (2008) 245–286
19. Carmagnola, F., Cena, F.: User identification for cross-system personalisation. Information Sciences: an International Journal **179**(1-2) (2009) 16–32
20. Abel, F., Henze, N., Herder, E., Krause, D.: Interweaving public user profiles on the web. In Bra, P.D., Kobsa, A., Chin, D.N., eds.: User Modeling, Adaptation, and Personalization, 18th International Conference, UMAP 2010, Big Island, HI, USA, June 20-24, 2010. Proceedings. Volume 6075 of Lecture Notes in Computer Science., Springer (2010) 16–27
21. Mehta, B.: Learning from what others know: Privacy preserving cross system personalization. In Conati, C., McCoy, K.F., Paliouras, G., eds.: User Modeling. Volume 4511 of Lecture Notes in Computer Science., Springer (2007) 57–66
22. Sakaki, T., Okazaki, M., Matsuo, Y.: Earthquake shakes twitter users: real-time event detection by social sensors. In: Proceedings of the 19th international conference on World wide web. WWW '10, New York, NY, USA, ACM (2010) 851–860
23. Lerman, K., Ghosh, R.: Information contagion: an empirical study of spread of news on digg and twitter social networks. In: Proceedings of 4th International Conference on Weblogs and Social Media (ICWSM). (May 2010)
24. Dong, A., Zhang, R., Kolari, P., Bai, J., Diaz, F., Chang, Y., Zheng, Z., Zha, H.: Time is of the essence: improving recency ranking using twitter data. In: WWW '10: Proceedings of the 19th international conference on World wide web, New York, NY, USA, ACM (2010) 331–340
25. Liu, J., Dolan, P., Pedersen, E.R.: Personalized news recommendation based on click behavior. In Rich, C., Yang, Q., Cavazza, M., Zhou, M.X., eds.: Proceeding of the 14th international conference on Intelligent user interfaces (IUI '10), New York, NY, USA, ACM (2010) 31–40
26. Passant, A., Hastrup, T., Bojars, U., Breslin, J.: Microblogging: A Semantic Web and Distributed Approach. In Bizer, C., Auer, S., Grimnes, G.A., Heath, T., eds.: Proceedings of the the 4th Workshop Scripting For the Semantic Web (SFSW '08) co-located with ESWC '08, Tenerife, Spain. Volume 368., CEUR-WS.org (2008)