Towards a Framework for Adaptive Faceted Search on Twitter

Ilknur Celik[1], Fabian Abel[1], Patrick Siehndel[2]

[1] Web Information Systems, Delft University of Technology
{celik,abel}@tudelft.nl
[2] L3S Research Center, Leibniz University Hannover, Germany
siehndel@l3s.de

Abstract. In the last few years, Twitter has become a powerful tool for publishing and discussing information. Yet, content exploration in Twitter requires substantial efforts and users often have to scan information streams by hand. In this paper, we approach this problem by means of faceted search. We propose strategies for inferring facets and facet values on Twitter by enriching the semantics of individual Twitter messages and present different methods, including personalized and context-adaptive methods, for making faceted search on Twitter more effective. We conduct a preliminary analysis that shows that semantic enrichment of tweets is essential for faceted search on Twitter and that there is essential need for adaptive faceted search on Twitter. Furthermore, we propose an evaluation methodology that allows us to automatically evaluate the quality of adaptive faceted search on Twitter without requiring expensive user studies.

Key words: faceted search, twitter, semantic enrichment, adaptation

1 Introduction

With the growing information space on the Web and the increasing popularity of Social Media, Social Web applications became part of daily activities as well as the source of information for millions of people. The dynamic nature of the Web and the diversity of the users along with the heavy information load demanded some form of adaptation or personalization in many Web-based applications in various domains. Nowadays, many Social Web applications are suffering from similar information overload problems, where the users of these applications find it difficult to read, find and follow the relevant and interesting information shared by a large network of other users. Our research focuses on tackling information overload in one of the most popular of these applications, Twitter.

Twitter is the most popular micro-blogging site and a growing Social Web phenomenon that is attracting interest from different types of people all around the world for a variety of different purposes, such as fast communication, work, status updates, following news, sports, events, opinions, hot topics, and so on [1–8]. With millions of Twitter messages (tweets) per day, highly active users are

estimated to receive hundreds of tweets every day[3]. Due to the lack of any adaptive or personalized navigation support in Twitter, users may get lost, become de-motivated and frustrated in this network of information overload [10]. Accessing required or interesting fresh content easily is vital in today's information age. Hence, there is a need for an effective personalized searching option from the users' point of view that would assist them in following the optimal path through a series of facets to find the information they are looking for, while providing a structured environment for relevant content exploring. Our research focuses on investigating ways to enhance searching and browsing in microblogging sites like Twitter by means of adaptive and personalized faceted search.

Searching and browsing are, indeed, somewhat limited in Twitter. For example, one can search for tweets by a keyword or by a user in a timeline that would return the most recent posts. So, if a user wants to see the different tweets about a field of sports, and were to search for "sports" in Twitter, only the recent tweets that contain the word "sports" would be listed to the user. Many tweets that do not contain the search keyword, but are about different sport events, sport games and sport news in general, would not be returned. Moreover, the Twitter keyword search differs from the general Web search due to the restricted message size of 140 characters in Twitter [9]. Traditional faceted search interfaces allow users to search for items by specifying queries regarding different dimensions and properties of the items (facets) [11]. For example, online stores such as eBay[4] or Amazon[5] enable narrowing down their users' search for products by specifying constraints regarding facets such as the price, the category or the producer of a product. In contrast, information on Twitter is rather unstructured and short, which does not explicitly feature facets. This puts constrains on the size and the number of keywords, as well as facets, that can be used as search parameters without risking to filter out many relevant results. Hence, searching by more than one topic (multiple facets), such as "sport events", would return only those recent tweets that contain both of these words and miss tweets like *"Off to BNP Paribas at Indian Wells"*, which mentions the name and the location of a sport event without necessarily including the keywords. In this paper, we introduce an adaptive faceted search framework for Twitter and investigate how to extract facets from tweets, how to design appropriate faceted search strategies on Twitter and how to evaluate such a framework. Our main contributions can be summarized as follows.

Semantic Enrichment We present methods for enriching the semantics of tweets by extracting facets (entities and topics) from tweets and related external Web resources.

User and Context Modeling Given the semantically enriched tweets, we propose user and context modeling strategies that identify (current) interests of a given Twitter user and allow for contextualizing the demands of this user.

[3] `http://techcrunch.com/2010/06/08/twitter-190-million-users/`
[4] `http://ebay.com/`
[5] `http://amazon.com/`

Adaptive Faceted Search We introduce faceted search strategies for content exploration on Twitter and propose methods that adapt to the interests and context of a user.

Evaluation Framework We present an evaluation environment based on simulated users to evaluate different strategies in our adaptive faceted search engine on Twitter.

2 Related Work and Our Motivation

The exponential growth of Twitter has attracted significant amount of research from various perspectives and fields recently. In this section, we focus on the related work that motivates and inspires our work, as well as relating our work to the existing literature.

2.1 Content Exploration on Twitter

A prototype for topic-based browsing in Twitter was proposed after observing how the users manage the incoming flood of updates [10]. This prototype interface, called Eddi, visualizes a user's Twitter feed using topic clusters constructed via a topic identification algorithm without using any semantics or natural language processing. This approach, however, does not find the relations between the topics or perform any recommendation of related topics. While it provides a means for browsing through a user's own feed by topics, our ambition is to infer relations between entities of all tweets in the network in order to adapt the list of facets presented to contain the related entities of the tweet of interest even outside of the user's feed. The aim is to provide a means where not only the users can easily reach to the information they are looking for by controlling their search parameters as they move along, but can also browse the related information about the current subject of interest by related people, countries, cities, events, and other selected facets.

2.2 Semantic Enrichment of Tweets

The main problem in searching microblogging platforms is the size of the messages. For example, the Twitter messages, with 140 characters limit, are too short to extract meaningful semantics on their own. Furthermore users tend to use abbreviations and short-form for words to save space, as well as colloquial expressions, which make it even harder to infer semantics from tweets. Rowe et al. mapped tweets to conference talks and exploited metadata of the corresponding research papers to enrich the semantics of tweets to better understand the semantics of the tweets published in conferences [12]. We follow a similar approach to this, except we try to enrich the tweets in general and not in a restricted domain like scientific conferences. A study by Kwak et al. revealed that the majority of the trending topics in Twitter are either headline or persistent news, with 85% of all the posted tweets being related to news, claiming Twitter is used more as a news media than a social network [4]. Consequently, we try to map tweets to news articles on the Web over the same time period in order to enrich them and to allow for extracting more entities to generate richer facets.

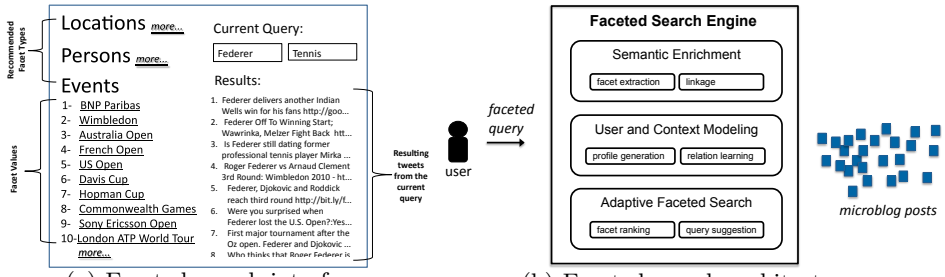
2.3 User and Context Modeling for Adaptive Faceted Search in Twitter

We also try to discover the relations between the extracted entities by studying different strategies in order to determine relatedness relations between entities such as persons related to an event and identify any temporal constraints on such relations. These learnt relations between entities can be utilized to ease the search by grouping together the related facets and recommending the most relevant facets that the user is looking for. Marinho et al. proposed a method for collabulary learning which takes a folksonomy and domain-expert ontology as input and performs semantic mapping to generate an enriched folksonomy [13]. An algorithm based on frequent itemsets techniques is then applied to learn an ontology over this enriched folksonomy. A similar approach exploited frequent itemsets to learn association rules from tagging activities [14]. We study the co-occurrence frequencies of entity pairs and compare these with other strategies for tweets in combination with news articles to learn relations between these entities.

In addition to adapting the facets to the current search, we aim at adapting the facet values to the current state of the users in order to personalize the search and content exploration. Liu et al. analyzed content-based recommenders for Google News and showed that interests in news topics such as technology, politics, et cetera change over time [15]. They also predicted user interests and showed that these user profiles in combination with recent trends on Google News outperform collaborative filtering. Similarly, Chen et al. studied content recommendation in Twitter and found out that both topic and relevance are important considerations [16]. They also observed that URLs extracted from the user's close social group is more successful than the most popular ones. Correspondingly, we observe the users' past activities to infer their recent interests based on their recent tweets and re-tweets. In other words, we build a profile of user interests in accordance with entities and topics, which is then used to adapt ranking of the facet values. Re-arranging the facet values according to user history and interests in line with the trendy topics can accelerate and thus improve the searching experience.

3 Faceted Search on Twitter

On Twitter, facets describe properties of a Twitter message. For example, persons that are mentioned in a tweet or events a tweet refers to. Oren et al. [11] formulate the problem of faceted search in RDF terminology. Given an RDF statement $(subject, predicate, object)$, the faceted search engine interprets (i) the subject as the actual resource that should be returned by the engine, (ii) the predicate as the facet type and (iii) the object as the facet value (restriction value). A faceted query (facet-value pair) that is sent to a faceted search engine thus consists of a predicate and an object. We follow this problem formulation proposed by Oren et al. [11] and interpret tweets as the actual resources the faceted search engine should return. If a tweet (subject) mentions an entity then



(a) Faceted search interface (b) Faceted search architecture

Fig. 1. Adaptive faceted search on Twitter: (a) example interface and (b) architecture of the faceted search engine.

the type of the entity is considered as facet type (predicate) and the actual identifier of the entity is considered as facet value (object). For example, given a tweet t that refers to the tennis player "Federer", the corresponding URI of the entity ($URI_{federer}$) and the URI of the entity type (URI_{person}) are used to describe the tweet by means of an RDF statement: $(t, URI_{person}, URI_{federer})$.

Figure 1(a) illustrates how we envision the corresponding faceted search interface that allows users to formulate faceted queries. Given a list of facet values which are grouped around facet types such as locations, persons and events, users can select facet-value pairs such as $(URI_{event}, URI_{wimbeldon})$ to refine their current query $((URI_{person}, URI_{federer}), (URI_{sportsgame}, URI_{tennis}))$. A faceted query thus may consist of several facet-value pairs. Only those tweets that match all facet-value constraints will be returned to the user. The ranking of the tweets that match a faceted query is a research problem of its own and could be solved by exploiting the popularity of tweets – e.g. measured via the number of retweets or via the popularity of the user who published the tweet (cf. [17]). The core challenge of the faceted search interface is to support the facet-value selection as good as possible. Hence, the facet-value pairs that are presented in the faceted search interface (see left in Figure 1(a)) have to be ranked so that users can quickly narrow down the search result lists until they find the tweets they are interested in. Therefore, the *facet ranking problem* can be defined as follows.

Definition 1 (Facet Ranking Problem). *Given the current query F_{query}, which is a set of facet-value pairs $(predicate, object) \in F_{query}$, the hit list H of resources that match the current query, a set of candidate facet-value pairs $(predicate, object) \in F$ and a user u, who is searching for a resource t via the faceted search interface, the core challenge of the faceted search engine is to rank the facet-value pairs F. Those pairs should appear at the top of the ranking that restrict the hit list H so that u can retrieve t with the least possible effort.*

The effort, which u has to invest to narrow down the search result list H, can be measured by click and scroll operations. Strategies for facet ranking are discussed in Section 3.2.

3.1 Architecture for Adaptive Faceted Search on Twitter

Figure 1(b) illustrates the architecture of the engine that we propose for faceted search on Twitter. The main components of the engine are the following.

Semantic Enrichment The semantic enrichment layer aims to extract facets from tweets and generate RDF statements that describe the facet-value pairs which are associated with a Twitter message. In particular, each tweet is processed to identify entities (facet values) that are mentioned in the message. We therefore make use of the OpenCalais API[6], which allows for the extraction of 39 different types of entities (facet types) including persons, organizations, countries, cities and events. As Twitter messages are limited to 140 characters, the extraction of entities from tweets is a non-trivial problem. Thus, we introduced a set of strategies that link tweets with external Web resources (news articles) and propagate the semantics extracted from these resources to the related tweets in [18]. For example, given a tweet "This is great http://bit.ly/2fRds1t", we extract entities from the referenced resource (http://bit.ly/2fRds1t) and attach the extracted entities to the tweet. In our analysis, we show that this semantic enrichment allows us to significantly better prepare the tweets for faceted search than enrichment which is merely based on tweets.

User and Context Modeling In order to adapt the facet ranking to the people who are using the faceted search engine, we propose user modeling and context modeling strategies. The user modeling strategies model the interests of the users in certain facet values (entities and topics). We therefore exploit the tweets that have been published (including re-tweets) by a user. In future work, we also plan to consider click-through data from the faceted search engine. Context modeling covers mining of new knowledge from the Twitter data. We therefore propose relation learning strategies that exploit co-occurrence of entities in Twitter messages to infer typed relationships between entities [19].

Adaptive Faceted Search Based on the semantically enriched tweets, the learnt relationships between entities extracted from tweets and the user profiles generated by the user modeling layer, the adaptive faceted search layer solves the actual facet ranking problem. It provides methods that adapt the facet-value pair ranking to the given context and user. Furthermore, it provides query suggestions by exploiting the relations learnt from the Twitter messages. Given the current facet query, which is a list of facet-value pairs where each value refers to an entity, we can exploit relationships between entities in order to identify entities that are related to those entities that occur in the current facet query. We leave the analysis of such query suggestions for future work. Instead, we focus on the facet ranking problem and propose different strategies for ranking facet-value pairs in the next subsection.

[6] `http://www.opencalais.com/`

3.2 Adaptive Faceted Search and Facet Ranking Strategies

Non-Personalized Facet Ranking A lightweight approach is to rank the facet-value pairs $(p, e) \in F$ based on their occurrence frequency in the current hit list H, the set of tweets that match the current query (cf. Definition 1):

$$rank_{frequency}((p, e), H) = |H_{(p,e)}| \tag{1}$$

$|H_{(p,e)}|$ is the number of (remaining) tweets that contain the facet-value pair (p, e) that can be applied to further filter the given hit list H. By ranking those facets that appear in most of the tweets, $rank_{frequency}$ minimizes the risk of filtering out relevant tweets but might increase the effort a user has to invest to narrow down search results.

Context-adaptive Facet Ranking The context-adaptive strategy exploits relationships between entities (facet values) to produce the facet ranking. A relationship is therefore defined as follows:

Definition 2 (Relationship). *Given two entities e_1 and e_2, a* relationship *between these entities is described via a tuple $rel(e_1, e_2, type, t_{start}, t_{end}, w)$, where* type *labels the relationship, t_{start} and t_{end} specify the temporal validity of the relationship and $w \in [0..1]$ is a weighting score that allows for specifying the strength of the relationship.*

The higher the weighting score w the stronger the relationship between e_1 and e_2. We use co-occurrence frequency as weighting scheme. Hence, given the enriched tweets, we count the number of tweets both entities (e_1 and e_2) are associated with. The context-adaptive facet ranking strategy ranks the facet-value pairs $(p, e) \in F$ according to $w(e_i, e)$, where e_i is a facet value that is already part of the given query: $(p_i, e_i) \in F_{query}$ (cf. Definition 1):

$$rank_{relation}((p, e), F_{query}) = \sum_i w(e_i, e) | (p, e_i) \in F_{query} \tag{2}$$

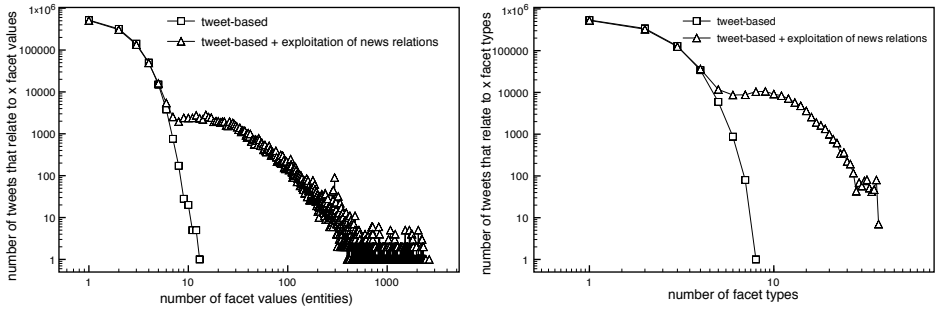
Hence, the context-sensitive strategy can only be applied in situations where the user has already made one selection, so that $|F_{query}| > 0$.

Personalized Facet Ranking The personalized facet ranking strategy adapts the facet ranking to a given user profile that is generated by the user modeling layer depicted in Figure 1(b). User profiles conform to the following model and specify a user's interest into a specific facet value (entity).

Definition 3 (User Profile). *The profile of a user $u \in U$ is a set of weighted entities where with respect to the given user u for an entity $e \in E$ its weight $w(u, e)$ is computed by a certain function w.*

$$P(u) = \{(e, w(u, e)) | e \in E, u \in U\}$$

Here, E and U denote the set of entities and users respectively.



(a) number of facet values per tweet (b) number of facet types per tweet

Fig. 2. Impact of semantic enrichment on (a) the number of facet values per tweet and (b) the number of distinct facet types per tweet.

Given the set of facet-value pairs $(p, e) \in F$ (see Definition 1), the personalized facet ranking strategy utilizes the weight $w(u, e)$ in $P(u)$ to rank the facet-value pairs:

$$rank_{personalized}((p, e), P(u)) = \begin{cases} w(u, e) \text{ if } w(u, e) \in P(u) \\ 0 \qquad \text{otherwise} \end{cases} \tag{3}$$

By combining the above three strategies it is possible to generate further facet ranking methods. A combination of two strategies can be realized by building the weighted average computed for a given facet-value pair (p, e) (e.g. $rank_{combined} = \alpha \cdot rank_{\alpha}((p, e)) + \beta \cdot rank_{\beta}((p, e))$).

4 Analysis of Faceted Search on Twitter

In our analysis, we study the characteristics of facets on Twitter. As described above, tweets do not feature many facets by nature. Therefore, strategies that enrich the semantic of tweets are required in order to derive facet-value pairs for tweets. In this section, we examine how the semantic enrichment supports the derivation of facets. Furthermore, we analyze the feasibility of the user and context modeling strategies for making faceted search on Twitter adaptive.

4.1 Analysis of Semantic Enrichment

As tweets do not provide facets related to the topic, our faceted search framework provides the functionality to enrich the semantics of tweets. To analyze the feasibility of our semantic enrichment component (see Section 3), we monitored the Twitter activities of more than 20,000 users over a period of more than two months and processed the data that we collected (1,671,389 tweets in total) to extract facet values from the tweets. For 62.91% of the tweets, we succeeded in extracting at least one entity that we can use as facet value. By making use of the semantic enrichment functionality that exploits links to external Web resources (and news articles in particular), we increased the coverage so that 66.77% of

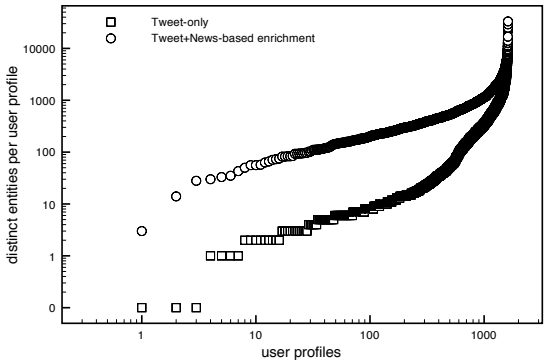


Fig. 3. Entity-based user profiles that can be exploited for personalized facet ranking.

the tweets which are enriched with facet values obtained from related news have at least one facet value. In the context of the news-based enrichment, we 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 comparing the timestamps. In previous work we showed that this method correlates news and tweets with an accuracy of more than 70% [20].

Figure 2(a) reveals that the number of facet values increases clearly when tweets are enriched with entities of related news articles. For example, less than 20 tweets exhibit more than 10 facet values in the case of semantic enrichment that is merely based on tweets . Given that tweets are limited to 140 characters, this observation is expected. Moreover, the number of different facet types per tweet also increases when linkage to news articles is exploited (see Figure 2(b)). In our current implementation, we differentiate between 39 different facet types, where persons, countries and organizations are the most popular types of facets. In Figure 2(b), we see that the tweet-based enrichment does not allow for more than 10 different types of facet types per tweets while the exploitation of news relations features more than 10,000 tweets that can be discovered via more than 10 different facet types, i.e. users can choose between various facets to narrow down the actual hit list (cf. Figure 1(a)).

4.2 Analysis of User and Context Modeling

The adaptation of the faceted search interface to the preferences of the user and therefore the personalized facet ranking strategy (see Equation 3) requires entity-based user profiles (see Definition 3). To analyze to what extent this method can succeed, we show the profile size of 1500 randomly selected user profiles in Figure 3. We see that the news-based enrichment results in profiles that provide more entities than the tweet-only based enrichment. For example, semantic enrichment based merely on tweets fails for three users as the size of the profile is zero for these users. In contrast, the news-based enrichment successfully generates profiles for all users. For more than 98% of the users, the number of distinct

entities per profile is even higher than 100. This indicates that news-based enrichment prevents from sparsity problems and thus allows for supporting the personalized facet ranking better than the tweets-only-based enrichment.

5 Evaluation Framework for Faceted Search

Evaluating the performance of faceted search is challenging. It usually requires query logs and click-through data, which is difficult to get for researchers, or calls for user studies, which are expensive if they are conducted on a large scale. In this section, we propose a novel technique for automatically evaluating the performance of faceted search on Twitter. Our evaluation methodology follows an idea introduced by Koren et al. [21] and exploits re-tweets as ground truth for estimating user relevance. The evaluation methodology is based on simulated users who behave in a predefined way. The utility of the interface is measured by the actions a simulated user needs to perform in order to find a relevant document.

General Setup. The general setup used for the evaluation process contains parameters describing the user interface itself and algorithms characterizing the simulated user behavior. In general, all faceted search user interfaces share some common characteristics and contains at least two parts: an area displaying the facets and a part showing the search results. For our evaluation process, the number of documents to be presented at a time, the number of different facets to be displayed and the number of elements which can be shown for each different facet need to be defined. We setup a basic framework for a search interface by defining these three parameters. Based on this interface, a user can perform different actions, where the goal is to find a relevant document. For every action we can define a cost, where the cost is related to the time a real user would need to accomplish this action. In our scenario a user can perform the following actions:

Select facet-value pair Basic action a user performs every time a facet-value pair is clicked, where the displayed search results are automatically updated after the selection (costs: 1).

View more facet-value pairs This action indicates that none of the currently displayed facet-value pairs are relevant for the user. By performing this action the user gets an additional amount of facet-value pairs related to one facet (costs: 2).

Show more documents This action allows the user to see more documents (tweets) matching the currently selected facet-values (costs: 2).

Select relevant tweet This action ends the current search (costs: 0).

Beside the actions mentioned above one could also consider the act of deselecting previously marked facet-values. In our search scenario, this action is not included as we assume that the users have perfect knowledge about the tweet they are looking for, and therefore a wrong selection will not take place.

Selection Strategies. The simulated users select facet-value pairs based on different strategies. The strategies we use for our evaluation are:

Random user This user randomly selects one of the displayed facet-values which matches the tweet he is looking for. If none of the displayed facet-value pairs matches the tweet, he randomly chooses one facet to see more facet-value pairs.

First-match user This user selects the first matching facet-value pair displayed by the interface. The basic idea behind this strategy is based on a user who directly clicks on a matching facet-value pair suggestion and do not look at all displayed facet value pairs to find the best matching one.

Greedy user This strategy tries to reduce the number of matching documents as fast as possible. This user selects the facet-value pair which occurs in the least number of remaining documents. This can be motivated by a user who selects the facet-value pair which is particularly important for the targeted tweet, in comparison to facet-value pairs which are related to many tweets.

Based on these facet selection strategies, the simulated user searches for a relevant document. The cost of this search is measured by the costs and number of actions a user needs to perform to find a relevant document.

Evaluation process. To measure the benefit of the proposed methods for faceted search, we evaluate the cost for a user to find relevant documents. Here, a tweet is relevant to a user, if the user re-tweeted this tweet. Re-tweeting a tweet indicates that the user has read the tweet and is to some extend interested in the content of the tweet. The proposed method is used to compare the costs of finding a relevant document when using the baseline ranking strategy based on frequency (non-personalized facet ranking) in comparison with context-adaptive facet ranking and personalized facet ranking.

6 Conclusions

In this paper, we presented an adaptive and personalized faceted search engine for Twitter, where we explained approaches for enriching the semantics of tweets, extracting facets, discovering relatedness information between entities and observing user activities to learn their behavior and interests in order to support users in their search for specific information or tweets. We proposed different strategies based on learnt relations together with user action history for adapting the search behavior as well as improving content exploration in Twitter. Furthermore, we introduced a generic evaluation environment based on Koren et al. [21] that will allow us to evaluate our strategies by simulated experiments, which constitutes part of our future research.

Acknowledgements The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no ICT 257831 (ImREAL project[7]).

[7] `http://imreal-project.eu`

References

1. Hughes, A.L., Palen, L.: Twitter Adoption and Use in Mass Convergence and Emergency Events. In: Proc. of ISCRAM. (2009)
2. Zhao, D., Rosson, M.B.: How and why people Twitter: the role that micro-blogging plays in informal communication at work. In: Proc. GROUP, ACM (2009) 243–252
3. Cha, M., Haddadi, H., Benevenuto, F., Gummadi, P.K.: Measuring User Influence in Twitter: The Million Follower Fallacy. In: Proc. of ICWSM, The AAAI Press (2010)
4. Kwak, H., Lee, C., Park, H., Moon, S.: What is twitter, a social network or a news media? In: Proc. of WWW, ACM (2010) 591–600
5. Lerman, K., Ghosh, R.: Information contagion: an empirical study of spread of news on digg and twitter social networks. In: Proc. of ICWSM, The AAAI Press (2010)
6. Java, A., Song, X., Finin, T., Tseng, B.: Why we twitter: understanding microblogging usage and communities. In: Proc. of WebKDD/SNA-KDD, ACM (2007) 56–65
7. Kaufman, S.J., Chen, J.: Where we Twitter. In: Proc. of Workshop on Microblogging: What and How Can We Learn From It? (2010)
8. Romero, D.M., Meeder, B., Kleinberg, J.: Differences in the mechanics of information diffusion across topics: Idioms, political hashtags, and complex contagion on twitter. In: Proc. of WWW, ACM (2011)
9. Teevan, J., Ramage, D., Morris,M.R.: #TwitterSearch: A Comparison of Microblog Search and Web Search. In: Proc. of WSDM, ACM (2011)
10. Bernstein, M., Kairam, S., Suh, B., Hong, L., Chi, E.H.: A torrent of tweets: managing information overload in online social streams. In: Proc. of Workshop on Microblogging: What and How Can We Learn From It? (2010)
11. Oren, E., Delbru, R., Decker, S.: Extending faceted navigation for rdf data. In: Proc. of ISWC, Springer (2006) 559–572
12. Rowe, M, Stankovic, M., Laublet, P.: Mapping Tweets to Conference Talks: A Goldmine for Semantics. In: Proc. of SDoW, colocated with ISWC, CEUR-WS.org (2010)
13. Balby Marinho, L., Buza, K., Schmidt-Thieme, L.: Folksonomy-based collabulary learning. In: Proc. of ISWC, Springer (2008) 261–276
14. Hotho, A., Jäschke, R., Schmitz, C., Stumme, G.: Emergent Semantics in BibSonomy. In: Informatik für Menschen. Volume 94(2) of LNI, GI (2006)
15. Liu, J., Dolan, P., Pedersen, E.R.: Personalized news recommendation based on click behavior. In: Proc. of IUI, ACM (2010) 31–40
16. Chen, J., Nairn, R., Nelson, L., Bernstein, M., Chi, E.: Short and tweet: experiments on recommending content from information streams. In: Proc. of CHI, ACM (2010) 1185–1194
17. Weng, J., Lim, E.P., Jiang, J., He, Q.: Twitterrank: finding topic-sensitive influential twitterers. In: Proc. of WSDM, ACM (2010) 261–270
18. Abel, F., Gao, Q., Houben, G.J., Tao, K.: Analyzing User Modeling on Twitter for Personalized News Recommendations. In: Proc. of UMAP, Springer (2011)
19. Celik, I., Abel, F.: Learning Semantic Relationships between Entities in Twitter. In: Proc. of ICWE, (2011)
20. Abel, F., Gao, Q., Houben, G.J., Tao, K.: Semantic Enrichment of Twitter Posts for User Profile Construction on the Social Web. In: ESWC, Springer (2011)
21. Koren, J., Zhang, Y., Liu, X.: Personalized interactive faceted search. In: Proc. of WWW, ACM (2008) 477–486