

Temporal Effects on Hashtag Reuse in Twitter: A Cognitive-Inspired Hashtag Recommendation Approach

Dominik Kowald
Graz University of
Technology & Know-Center
Inffeldgasse 13/5
Graz, Austria
dkowald@know-center.at

Subhash Chandra Pujari
Graz University of
Technology
Inffeldgasse 13/5
Graz, Austria
s.pujari@student.tugraz.at

Elisabeth Lex
Graz University of
Technology
Inffeldgasse 13/5
Graz, Austria
elisabeth.lex@tugraz.at

ABSTRACT

Hashtags have become a powerful tool in social platforms such as Twitter to categorize and search for content, and to spread short messages across members of the social network. In this paper, we study temporal hashtag usage practices in Twitter with the aim of designing a cognitive-inspired hashtag recommendation algorithm we call $BLL_{I,S}$. Our main idea is to incorporate the effect of time on (i) individual hashtag reuse (i.e., reusing own hashtags), and (ii) social hashtag reuse (i.e., reusing hashtags, which has been previously used by a followee) into a predictive model. For this, we turn to the Base-Level Learning (BLL) equation from the cognitive architecture ACT-R, which accounts for the time-dependent decay of item exposure in human memory. We validate $BLL_{I,S}$ using two crawled Twitter datasets in two evaluation scenarios. Firstly, only temporal usage patterns of past hashtag assignments are utilized and secondly, these patterns are combined with a content-based analysis of the current tweet. In both evaluation scenarios, we find not only that temporal effects play an important role for both individual and social hashtag reuse but also that our $BLL_{I,S}$ approach provides significantly better prediction accuracy and ranking results than current state-of-the-art hashtag recommendation methods.

Keywords

Twitter; Hashtags; BLL Equation; ACT-R; TF-IDF; Hashtag Recommendation; Hashtag Reuse Prediction; Temporal Dynamics; Hashtag Usage Recency; Recommender Systems

1. INTRODUCTION

Over the past years, the microblogging platform Twitter has become one of the most popular social networks on the Web with more than 310 million monthly active users. Users can build a network of follower connections to other Twitter users, which means that they can subscribe to content posted by their *followees* [31, 23]. Twitter was also the

first social platform that adopted the concept of *hashtags*, as suggested by Chris Messina in 2007¹.

Hashtags are freely-chosen keywords starting with the hash character “#” to annotate, categorize and contextualize Twitter posts (i.e., tweets) [34, 12]. The advantage of hashtags is that anyone with an interest in a hashtag can track it and search for it [38], thus receiving content posted by somebody outside of their own Twitter network. For example, users can retrieve tweets created during the European football championship by searching for the hashtag *#euro2016*, even if they do not have a social link to the tweet producers. Meanwhile, many social platforms, such as Instagram and Facebook, have adopted hashtags as well.

Problem. Unsurprisingly, the widespread acceptance of hashtags has sparked a lot of research in the field of *hashtag recommendations* (see Section 6 for a selection of approaches) to support users in assigning the most descriptive hashtags to their posts. Existing methods typically utilize collaborative, content and topic features of tweets to recommend hashtags to users. Undoubtedly, these features play an important role in recommending hashtags that best describe a tweet. In this paper, however, we are especially interested in predicting which hashtags a user will likely apply in a newly created tweet given previous hashtag assignments.

The main problem we want to address is whether we can identify *temporal usage patterns* that influence if a Twitter user will likely utilize a certain hashtag in a tweet, given the hashtags she and/or her followees have used in the past. Our goal is to describe such temporal usage patterns by utilizing a model from human memory theory and to design a hashtag recommendation algorithm based on that. To the best of our knowledge, so far, few studies (e.g., [11]) have investigated the way temporal effects can be exploited in the hashtag recommendation process.

Approach and methods. We propose a cognitive-inspired hashtag recommendation algorithm we call $BLL_{I,S}$ that is based on temporal usage patterns of hashtags derived from empirical evidence. In essence, these patterns reflect how a person’s own hashtags as well as hashtags from the social network are utilized and reused. In our approach, we utilize the Base-Level Learning (BLL) equation from the cognitive architecture ACT-R [2, 3] to model temporal usage practices of hashtags. The BLL equation accounts for the time-dependent decay of item exposure in human memory. It quantifies the usefulness of a piece of information (e.g., a

©2017 International World Wide Web Conference Committee (IW3C2), published under Creative Commons CC BY 4.0 License.
WWW 2017, April 3–7, 2017, Perth, Australia.
ACM 978-1-4503-4913-0/17/04.
<http://dx.doi.org/10.1145/3038912.3052605>



¹<https://twitter.com/chris Messina/status/223115412>

hashtag) based on how frequently and how recently it was used by a user in the past and models this time-dependent decay by means of a power-law distribution. Thus, $BLL_{I,S}$ takes into consideration the frequency and recency of hashtags used by a user and her followees in the past.

We presented the BLL equation in our previous work as a model to recommend tags in social bookmarking systems such as BibSonomy and CiteULike [20, 19]. In the present work, we build upon these results by adopting the BLL equation to model the effect of time on the reuse of individual and social hashtags to build our hashtag recommendation algorithm. We demonstrate the efficacy of our approach in two empirical social networks crawled from Twitter. The first social network, termed *CompSci* dataset, is built upon the tweets of a sample of Twitter users, who have been identified as computer scientists in previous related work [10], and their followees. The second network, termed *Random* dataset, is built upon the tweets of a set of randomly chosen Twitter users and their followees. We experiment with these datasets to investigate the performance of our hashtag recommendation approach in two settings: (i) tweets of a domain-specific Twitter network, and (ii) tweets of a random network of Twitter users.

Contributions and findings. The main contributions of our work are two-fold. Firstly, our paper shows that time has a large effect on individual as well as social hashtag reuse in Twitter. Specifically, we observe a time-dependent decay of individual and social hashtag reuse that follows a power-law distribution. This finding paves the way for our idea to utilize the BLL equation as a predictive model to recommend hashtags for new tweets. Thus, our second contribution is that we design, develop and evaluate a personalized hashtag recommendation algorithm based on the BLL equation that outperforms current state-of-the-art approaches.

We implement the BLL equation in two variants, where the first one (i.e., $BLL_{I,S}$) predicts the hashtags of a user solely based on past hashtag usage, and the second one (i.e., $BLL_{I,S,C}$) combines $BLL_{I,S}$ with a content-based tweet analysis to also incorporate the text of the currently proposed tweet of a user. We evaluate our approach using standard evaluation protocols and metrics, and we find that both variants of our approach provide significantly higher prediction accuracy and ranking estimates than current state-of-the-art hashtag recommendation algorithms in both scenarios. We attribute this to the fact that our approach, in contrast to other related methods, mimics the way humans use and adapt hashtags by building upon insights from human memory theory (i.e., the BLL equation).

Structure of this paper. In Section 2, we continue by describing the crawling procedure of our two Twitter datasets and analyzing hashtag usage types in these datasets. Then, in Section 3, we study temporal usage patterns of individual and social hashtag reuse. In Section 4, we describe two variants of our approach (i.e., without and with the current tweet). This is followed in Section 5 by our evaluation methodology and experimental results. Finally, we discuss related work in the field in Section 6 and we give a summary of our findings as well as our future plans in Section 7.

2. DATASETS

In this section, we describe the data collection procedure and the two datasets we use for our study. Additionally, we

Dataset	$ U_S $	$ U $	$ T $	$ HT $	$ HTAS $
<i>CompSci</i>	2,551	91,776	5,649,359	1,081,403	9,161,842
<i>Random</i>	3,466	127,112	8,157,702	1,507,773	13,628,750

Table 1: Statistics of our *CompSci* and *Random* Twitter datasets. In this table, $|U_S|$ is the number of seed users, $|U|$ is the number of distinct users, $|T|$ is the number of tweets, $|HT|$ is the number of distinct hashtags and $|HTAS|$ is the number of hashtag assignments.

investigate individual as well as social hashtag reuse patterns in our datasets as a prerequisite for our hashtag recommendation approach.

Crawling strategy and dataset statistics. In order to address our research goals, we crawl two datasets using the Search API of Twitter². The final statistics of these datasets are illustrated in Table 1.

The first one (i.e., *CompSci* dataset) consists of researchers from the field of computer science and their followees, while the second one (i.e., *Random* dataset) consists of random people and their followees. Our idea is to test our hashtag recommendation approach in two different network settings: (i) a domain-specific one, in our case the domain of computer scientists, and (ii) a more general one consisting of random Twitter users. Our crawling strategy for both datasets comprises of the following four steps:

(a) **Crawl seed users.** We start with identifying and crawling a list of seed users U_S for each dataset. In the case of the *CompSci* dataset, we take the users who were identified as computer scientists in the work of [10]. In the case of the *Random* dataset, we use the Streaming API of Twitter³ to get a stream of tweets and extracted the user-ids to get our list of random seed users. From both user lists, we remove all users with more than 180 followees, which results in $|U_S| = 2,551$ seed users for the *CompSci* dataset and $|U_S| = 3,466$ seed users for the *Random* dataset. The threshold of using a maximum of 180 followees is chosen because the Twitter Search API only allows 180 requests to be made every 15 minutes, which gives us the possibility to crawl the tweets of all followees of a seed user within this reasonable time window.

(b) **Crawl followees.** Next, we use these follower relationships to crawl the followees F of the seed users in order to create a directed user network for analyzing the social influence on hashtag reuse. Based on the number of seed users, the average number of followees per seed user $|F|/|U_S| = 94$ in the case of the *CompSci* dataset and 72 in the case of the *Random* dataset. Following these notations, the set of followees of user u is denoted as F_u in the remainder of this paper. Overall, our crawling procedure gives us $|U| = 91,776$ distinct users for the *CompSci* dataset and $|U| = 127,112$ distinct users for the *Random* dataset.

(c) **Crawl tweets.** In the third step, we crawl the 200 most recent tweets of all the users and remove the tweets in which no hashtags are used. The threshold of a maximum of 200 most recent tweets is set because of another restriction of the Twitter Search API that only allows 200 tweets to be received per a single request. This crawling procedure results in $|T| = 5,649,359$ tweets for the *CompSci* dataset

²<https://dev.twitter.com/rest/public/search>

³<https://dev.twitter.com/streaming/overview>

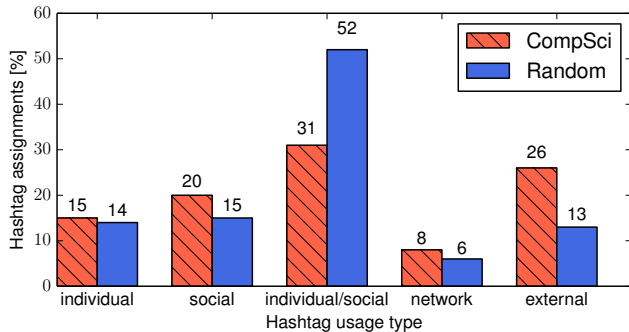


Figure 1: Analysis of hashtag usage types in our two datasets. For each hashtag assignment, we study whether the corresponding hashtag has been used by the same user before in time (“individual”), by some of the users she follows (“social”), by both (“individual/social”), by anyone else in the dataset (“network”) or neither of them (“external”). We find that between 66% and 81% of hashtag assignments in both datasets can be explained by individual or social hashtag usage (i.e., the sum of “individual”, “social” and “individual/social”).

with an average number of tweets per user $|T|/|U| = 61$, and $|T| = 8,157,702$ tweets for the *Random* dataset with $|T|/|U| = 64$. Our crawled tweets cover a time range from 2007 to 2015 in case of both datasets.

(d) Extract hashtag assignments. In the final step of our crawling procedure, we extract the hashtag assignments by searching for all words that start with a “#” character. This results in $|HTAS| = 9,161,842$ hashtag assignments for $|HT| = 1,081,403$ distinct hashtags in the *CompSci* network and $|HTAS| = 13,628,750$ for $|HT| = 1,507,773$ in the *Random* network. Thus, in both datasets, each distinct hashtag is used approximately 9 times on average and each user uses approximately 100 hashtag assignments in her tweets on average. Examples for popular hashtags are *#bigdata*, *#iot* and *#ux* in case of the *CompSci* dataset, and *#shahbag*, *#ff* and *#art* in case of the *Random* dataset.

Analysis of hashtag usage types. In our datasets, we analyze hashtag assignments as well as hashtag reuse practices with the aim of identifying the different types of hashtag usages as a prerequisite for our recommendation approach. Specifically, for each hashtag assignment, we study whether the corresponding hashtag has either been used by the same user before (“individual”), by some of her followees (“social”), by both (“individual/social”), by anyone else in the dataset (“network”) or by neither of them (“external”).

The results of this study are shown in Figure 1. We find that 66% of hashtag assignments in the *CompSci* dataset and 81% in the *Random* dataset can be explained by individual or social hashtag reuse. This finding further supports our choice to utilize these two types of influences (i.e., individual and social) to create our model. In contrast to these large numbers, the 6% to 8% of hashtags in the “network” category is relatively small. Interestingly, the amount of “external” hashtags is twice as high in the *CompSci* dataset (i.e., 26%) as in the *Random* one (i.e., 13%). Thus, in our datasets, computer scientists tend to use more hashtags, which have not been previously introduced in the network, than

random Twitter users. Because of this, we believe that the recommendation accuracy results would generally be lower in the *CompSci* dataset than in the *Random* one, which will be evaluated in Section 5. Summing up, both individual and social hashtags have an impact on the users’ choice of hashtags for a new tweet.

3. TEMPORAL EFFECTS ON HASHTAG REUSE IN TWITTER

In this section, we study to what extent temporal effects play a role in the reuse of individual and social hashtags in our two datasets (i.e., *CompSci* and *Random*). Specifically, we analyze the recency of hashtags assignments (i.e., the time since the last hashtag usage/exposure), as well as whether this effect of time-dependent decay follows a power-law or exponential distribution.

Temporal effects on individual hashtag reuse. The effect of time on individual hashtag reuse is visualized in the plots (a) and (b) of Figure 2. To put the x-scale of these plots onto a meaningful range, we set the threshold for the maximum hashtag reuse recency to one year (i.e., 8,760 hours). The plots show the individual hashtag reuse count plotted over the reuse recency of a hashtag *ht* by a user *u* in hours. Hence, for each hashtag assignment of a hashtag *ht* by user *u*, we take the time since the last usage of *ht* by *u* (i.e., the reuse recency) and pool together all hashtag assignments with the same recency value (i.e., the same time difference in hours). The individual reuse count for this recency value is then given by the size of the set of these hashtag assignments.

The two plots show similar results for both datasets and indicate that the more recently a hashtag *ht* was used by a user *u* in the past, the higher its individual reuse count is. Interestingly, there is a clear peak after 24 hours in both datasets, which further indicates that users typically use the same set of hashtags in this time span and thus, tend to tweet about similar topics on a daily basis. Furthermore, we also observe high R^2 values of nearly .9 for the linear fits in the log-log scaled plots, which indicates that a large amount of our data can be explained by a power function. This is also suggested by the power-law-based model of the BLL equation [3, 2]. In contrast, the linear fits in log-linear scaled plots only provide R^2 values of approximately .7, where high values would speak in favor of an exponential function.

Temporal effects on social hashtag reuse. Plots (c) and (d) of Figure 2 show the effect of time on the social hashtag reuse for the *CompSci* and *Random* datasets. These plots are created similarly as plots (a) and (b) but this time, we plot the social hashtag reuse count over the reuse recency of a hashtag *ht* by the followees F_u of user *u*. Hence, for each hashtag assignment of *ht* by *u*, we take the most recent usage timestamp of *ht* by F_u . The difference between this timestamp and the timestamp of the currently analyzed hashtag assignment indicates the time since the last social exposure of *ht* to *u*. Again, we set the threshold for the maximum hashtag reuse recency to one year (i.e., 8,760 hours).

In these plots, we observe similar results for the two datasets since, in both cases, the more recently a user was exposed to a hashtag, the higher its social reuse count is. Furthermore, there is again (i) a clear peak after 24 hours, and (ii) the R^2 values for the linear fits in the log-log scaled plots

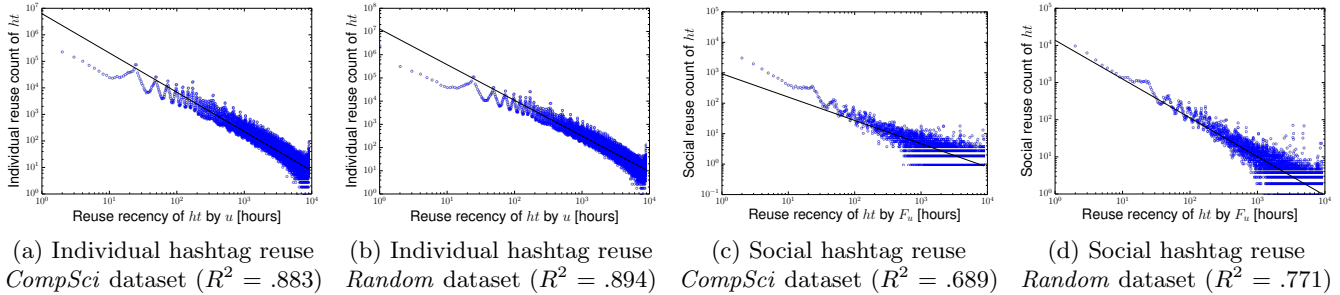


Figure 2: The effect of time on individual and social hashtag reuse for the *CompSci* and *Random* datasets (plots are in log-log scale). Plots (a) and (b) show that the more recently a hashtag ht was used by a user u , the higher its individual reuse count (i.e., people tend to reuse hashtags that have been used very recently by their own). Plots (c) and (d) show that the more recently a user u was exposed to a hashtag ht , which was used by her followees F_u , the higher its social reuse count (i.e., people tend to reuse hashtags that have been used recently in the social network). Additionally, we report the R^2 estimates for the linear fits of the data. We find that temporal effects play an important role in individual and social hashtag reuse in both datasets.

Dataset	Parameter	Individual ht reuse	Social ht reuse
<i>CompSci</i>	x_{min}	141	1
	α	1.699	1.242
	R	188	164
<i>Random</i>	x_{min}	141	1
	α	1.723	1.269
	R	235	294

Table 2: Power-law vs. exponential time-dependent decay. We see that a power function provides a better fit than an exponential function ($R > 0$ with $p < .001$) for explaining temporal effects on individual and social hashtag reuse in our two datasets.

(i.e., $R^2 = .7$) are larger than in the log-linear scaled plots (i.e., $R^2 = .4$), which speaks in favor of a power function. We now study if this is really the case.

Power-law vs. exponential time-dependent decay. The question whether a power or an exponential function is better suited to model the time-dependent decay of hashtag reuse is of interest especially for the design of our hashtag recommendation approach since both types of functions have been used in the area of time-aware recommender systems. While the BLL equation suggests the use of a power function to model the decay of item exposure in human memory [3], related hashtag recommender approaches, such as the one proposed in [11], use an exponential function for this purpose. As already mentioned, the visual inspection of Figure 2 and the R^2 values of the linear fits favor a power function. However, [5] has shown that this least squares-based method can lead to misinterpretations and thus, a likelihood ratio-based test is suggested.

We use the Python implementation [1] of the method described in [5] to validate if a power function produces a better fit than an exponential one. The results of this test are shown in Table 2. The main value of interest here is the log-likelihood ratio R between the two functions. As we see, $R > 0$ in all four cases with $p < .001$. This means that the power function indeed provides a better fit than the exponential function for explaining temporal effects on individual and social hashtag reuse. We also provide the x_{min} and α values of the fits. In this respect, the α slopes can be used to

set the d parameter of the BLL equation (i.e., $d = 1.7$ in the individual case and 1.25 in the social case, see Section 4). Interestingly, these values are much higher than the suggested value of BLL’s d parameter, which is $.5$ [2]. We believe that this is the case because tweeting is more strongly influenced by temporal interest drifts than other applications studied in the ACT-R community (e.g., [3]).

Finding 1: Temporal effects have an important influence on both individual as well as social hashtag reuse. People tend to reuse hashtags that were used very recently by their own and/or by their Twitter followees. Furthermore, a power function is better suited to model this time-dependent decay than an exponential one. This suggests that the BLL equation from the cognitive architecture ACT-R should be a suitable model for designing our time-dependent hashtag recommendation algorithm.

4. A COGNITIVE-INSPIRED HASHTAG RECOMMENDATION APPROACH

In the previous section, we have shown that temporal effects are important factors when users reuse individual and social hashtags. In this section, we use these insights as a basis to design our hashtag recommendation approach illustrated in Figure 3. Thus, we distinguish between hashtag recommendations without (*Scenario 1*) and with (*Scenario 2*) incorporating the current tweet t .

Whereas the first variant of our approach solely uses the past hashtags of a user u and/or her followees F_u , the second variant also utilizes the content of the current tweet t . Hence, these two scenarios also differ in their possible use cases since the first one aims to foresee the topics a specific user will tweet about based on the predicted hashtags, whereas the second one aims to support a user in finding the most descriptive hashtags for a new tweet text [9].

Source code. For reasons of reproducibility, we implement and evaluate our approach by extending our open-source tag recommender benchmarking framework *TagRec*. The source code and framework is freely accessible for scientific purposes via our Github repository⁴.

⁴<https://github.com/learning-layers/TagRec>

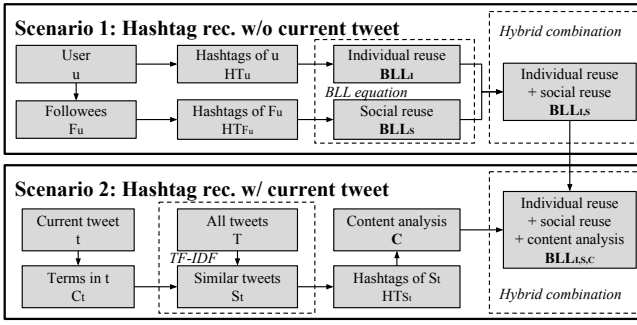


Figure 3: Schematic illustration of our cognitive-inspired approach for hashtag recommendations. We implement our approach in two scenarios (i.e., without and with incorporating the content of the current tweet). In *Scenario 1*, we use the BLL equation to realize (i) the individual BLL_I algorithm, (ii) the social BLL_S algorithm, and (iii) the hybrid $BLL_{I,S}$ algorithm, which combines both. In *Scenario 2*, we use a content analysis to identify similar tweets for a currently proposed tweet t and identify the hashtags of the most similar ones. We combine this content-based tweet analysis with our $BLL_{I,S}$ method to provide personalized and content-aware hashtag recommendations in the form of our hybrid $BLL_{I,S,C}$ approach.

4.1 Scenario 1: Hashtag rec. w/o current tweet

For the first variant of our approach, we ignore the content of the current tweet t and solely utilize past hashtag usages. As already stated, we use the BLL equation coming from the cognitive architecture ACT-R [2, 3] for this task. We go for a cognitive-inspired approach, since we know from research on the underlying mechanisms of social tagging that the way users choose tags for annotating resources (e.g., Web links) strongly corresponds to processes in human memory and its cognitive structures [6, 36]. The BLL equation quantifies the general usefulness of a piece of information (e.g., a word or hashtag) by considering how frequently and recently it was used by a user in the past. Formally, it is given by:

$$B_i = \ln\left(\sum_{j=1}^n t_j^{-d}\right) \quad (1)$$

where B_i is the base-level activation of a memory unit i and n is the frequency of i 's occurrences in the past (i.e., how often i was used by u). Furthermore, t_j states the recency (i.e., the time since the j th occurrence of i) and the exponent d accounts for the power-law of time-dependent decay. As visualized in *Scenario 1* of Figure 3, we adopt the BLL equation for (i) modeling the reuse of individual hashtags (BLL_I), (ii) modeling the reuse of social hashtags (BLL_S), and (iii) combining the former two into a hybrid recommendation approach ($BLL_{I,S}$).

Modeling individual hashtag reuse. In order to model the reuse of individual hashtags, we define the individual base-level activation $B_I(ht, u)$ of a hashtag ht for a user u as follows:

$$B_I(ht, u) = \ln\left(\sum_{j=1}^n (TS_{ref} - TS_{ht,u,j})^{-d_I}\right) \quad (2)$$

where n denotes the number of times ht was used by u in the past (i.e., $|HTAS_{ht,u}|$) and the term $TS_{ref} - TS_{ht,u,j}$ states the recency of the j th usage of ht by u . In this respect, TS_{ref} is the reference timestamp (i.e., when recommendations should be calculated) and $TS_{ht,u,j}$ is the timestamp when ht was used by u for the j th time. Based on the results of our analysis presented in Table 2, we set the individual time-dependent decay factor d_I to 1.7.

Modeling social hashtag reuse. We model the reuse of social hashtags in a similar way but instead of analyzing how frequently and recently a hashtag ht was used by user u , we analyze how frequently and recently ht was used by the set of followees F_u of u . Thus, we formulate the social base-level activation $B_S(ht, u)$ of ht for u as follows:

$$B_S(ht, u) = \ln\left(\sum_{j=1}^m (TS_{ref} - TS_{ht,F_u,j})^{-d_S}\right) \quad (3)$$

where m is the number of times ht was used by F_u before the reference timestamp TS_{ref} (i.e., $|HTAS_{ht,F_u}|$). The term $TS_{ref} - TS_{ht,F_u,j}$ states the recency of the j th exposure of ht to u caused by F_u , where $TS_{ht,F_u,j}$ is the timestamp when ht was used by F_u for the j th time. As when modeling the individual hashtag reuse, we set the social time-dependent decay factor d_S based on the results of our analysis in Table 2 (i.e., $d_S = 1.25$).

Combining individual and social hashtag reuse. As we have formalized the individual as well as social hashtag reuse, we want to mix both components in form of a hybrid approach using a linear combination [15]. Hence, in order to be able to add the individual and social base-level activations $B_I(ht, u)$ and $B_S(ht, u)$, we have to map these values onto a common range of 0 to 1 that add up to 1. Therefore, we define the softmax functions $\sigma(B_I(ht, u))$ and $\sigma(B_S(ht, u))$ as proposed by [30, 20]. This is given by:

$$\sigma(B_I(ht, u)) = \frac{\exp(B_I(ht, u))}{\sum_{ht' \in HT_u} \exp(B_I(ht', u))} \quad (4)$$

where HT_u is the set of distinct hashtags used by u . For $B_S(ht, u)$, the softmax function $\sigma(B_S(ht, u))$ can be calculated in the same way but on the basis of HT_{F_u} (i.e., the set of hashtags used by u 's followees F_u). Taken together, the combined base-level activation $B_{I,S}$ for our $BLL_{I,S}$ approach is given by a linear combination:

$$B_{I,S}(ht, u) = \beta \underbrace{\sigma(B_I(ht, u))}_{BLL_I} + (1 - \beta) \underbrace{\sigma(B_S(ht, u))}_{BLL_S} \quad (5)$$

where the β parameter can be used to give weights to the two components. Based on experimentation, we set β to .5 to equally weigh the individual and social influence. As indicated in Equation 5 and Figure 3, we can also calculate predictions either solely based on the individual hashtag reuse, referred as BLL_I , or solely based on the social hashtag reuse, referred as BLL_S .

4.2 Scenario 2: Hashtag rec. w/ current tweet

As shown in *Scenario 2* of Figure 3, the second variant of our approach aims to provide hashtag suggestions while also incorporating the content of the currently proposed tweet t . Thus, we build on the unpersonalized method proposed by [42] to find hashtags of similar tweets and combine this

method with our $BLL_{I,S}$ approach to generate personalized and content-aware recommendations.

Content-based tweet analysis. We analyze the content of tweets in order to find similar tweets for a target tweet t and to extract the hashtags of these similar ones. Therefore, we incorporate the term frequency-inverse document frequency (TF-IDF) statistic, which identifies the importance of a term for a document in a collection of documents. TF-IDF can be further used to calculate the similarity between two documents d and \bar{d} by summing up the TF-IDF scores of d 's terms in \bar{d} . When applying this statistic to Twitter, we treat tweets as documents and calculate the similarity between the target tweet t and a candidate tweet \bar{t} as follows:

$$sim(t, \bar{t}) = \sum_{c \in C_t} n_{c, \bar{t}} \times \log\left(\frac{|T|}{|\{t' : c \in t'\}|}\right) \quad (6)$$

where C_t are the terms in the text of target tweet t , $n_{c, \bar{t}}$ is the number of times $c \in C_t$ occurs in the candidate tweet \bar{t} , $|T|$ is the number of tweets in the dataset and $|\{t' : c \in t'\}|$ is the number of times c occurs in any tweet $t' \in T$. The first factor of this equation reflects the term frequency TF , whereas the second factor of this equation reflects the inverse document frequency IDF [42].

Based on these similarity values, we identify the most similar tweets S_t for t and extract the hashtags used in these tweets (i.e., HT_{S_t}). For each hashtag $ht \in HT_{S_t}$, we assign a content-based score $CB(ht, t)$, which is the highest similarity value within the most similar tweets S_t in which ht occurs. We implement this method using the Lucene-based full-text search engine Apache Solr 4.7.10⁵. Based on Solr's software documentation and our own experimentation, we set the minimum term frequency tf to 2 and the minimum document frequency df to 5.

Combining personalized and content-aware hashtag rec. We combine our personalized $BLL_{I,S}$ approach with this content-based analysis (C) in order to generate personalized hashtag recommendations (see Figure 3). Again, we achieve this via a linear combination of both approaches. Taken together, the top- k recommended hashtags $\widetilde{HT}_{u,t}$ for u and t are given by a linear combination:

$$\widetilde{HT}_{u,t} = \arg \max_{ht \in HT_{u,t}} \left(\lambda \underbrace{B_{I,S}(ht, u)}_{BLL_{I,S}} + (1 - \lambda) \underbrace{\sigma(CB(ht, t))}_C \right) \quad (7)$$

where $\widetilde{HT}_{u,t}$ is the set of candidate hashtags for u and t (i.e., $HT_u \cup HT_{F_u} \cup HT_{S_t}$). The λ parameter is used to give weights to the personalized and content-aware components. To that end, we set λ to .3 based on experimentation. Please note that the content-based score $CB(ht, t)$ has to be normalized using the softmax function (see Equation 4), whereas $B_{I,S}(ht, u)$ is already normalized (see Equation 5). This finally constitutes our personalized hashtag recommendation algorithm termed $BLL_{I,S,C}$.

5. EVALUATION

In this section, we present the evaluation of our approach. This includes the methodology used as well as the results in terms of recommendation accuracy and ranking for our two scenarios (i.e., with and without the current tweet).

⁵<http://lucene.apache.org/solr/>

5.1 Methodology

The methodology of our evaluation is given by the evaluation protocol, evaluation metrics and baseline algorithms used in our study.

Evaluation protocol. In order to split our datasets into training and test sets, we use an established leave-one-out evaluation protocol from research on information retrieval and recommender systems [15]. Therefore, for each seed user in our datasets (see Section 2) with at least two tweets (i.e., 2,020 users in the *CompSci* dataset and 2,679 users in the *Random* dataset), we determine her most recent tweet and put it (and its hashtags) into the test set. The remaining tweets are then put into the training set. This protocol ensures not only that the hashtags of at least one tweet per user are available for training but also that the chronological order of the data is preserved (i.e., future hashtags are predicted based on usage patterns of past ones). We use these sets in the following two evaluation scenarios:

(a) *Hashtag rec. w/o current tweet.* In the first scenario, we ignore the content of the currently proposed tweet (i.e., the one in the test set) and solely provide hashtag predictions based on the current user-id. Thus, in *Scenario 1*, we are able to evaluate all test set tweets.

(b) *Hashtag rec. w/ current tweet.* In the second scenario, we also incorporate the content of the current tweet. In this setting, we only evaluate the test set entries, which do not include retweets (i.e., 954 test set tweets in the *CompSci* dataset and 1,504 test set tweets in the *Random* dataset). The reason for excluding the retweets from the test set in *Scenario 2* is that searching for similar tweets in the training set would result in identical tweets with identical hashtags, which would heavily bias our evaluation (see also [42]).

Evaluation metrics. To finally quantify the quality of the hashtag recommendation algorithms, for each test set entry, we compare the top-10 hashtags an algorithm predicts for the given user u and tweet t (i.e., $\widetilde{HT}_{u,t}$) with the set of relevant hashtags actually used by user u in tweet t .

This comparison is done using various evaluation metrics known from the field of recommender systems. Specifically, we report Precision (P) and Recall (R) for $k = 1$ to 10 predicted hashtags by means of Precision/Recall plots, and F1-score (F1@5) for $k = 5$ predicted hashtags. We set $k = 5$ for the F1-score since F1@5 was also used as the main evaluation metric in the well-known ECML PKDD 2009 discovery challenge⁶. Additionally, we report three ranking-dependent metrics, namely (i) Mean Reciprocal Rank (MRR@10), (ii) Mean Average Precision (MAP@10), and (iii) Normalized Discounted Cumulative Gain (nDCG@10) for $k = 10$ predicted hashtags [13].

Baseline algorithms. We compare our approach to a rich set of 9 state-of-the-art hashtag recommendation algorithms:

(a) *MP_I.* The *Most Popular Individual Hashtags* algorithm ranks the hashtags based on the frequency in the hashtag assignments of current user u . MP_I is also referred to as Most Popular Tags by User (MP_u) in tag recommendation literature [15].

(b) *MR_I.* *Most Recent Individual Hashtags* is a time-dependent variant of MP_I . MR_I suggests the k most recently used hashtags of current user u [4]. Our BLL_I ap-

⁶<http://www.kde.cs.uni-kassel.de/ws/dc09/evaluation>.

Dataset	Metric	Scenario 1: Hashtag rec. w/o current tweet										Scenario 2: Hashtag rec. w/ current tweet		
		MP _I	MR _I	BLL _I	MP _S	MR _S	BLL _S	MP	FR	CF	BLL _{I,S}	SR	TCI	BLL _{I,S,C}
CompSci	F1@5	.086	.098	.101	.022	.076	.118	.006	.083	.099	.153***	.139	.182	.200*
	MRR@10	.136	.188	.193	.032	.122	.187	.007	.130	.163	.268***	.264	.334	.395***
	MAP@10	.143	.195	.202	.033	.128	.205	.007	.136	.169	.285***	.283	.354	.417***
	nDCG@10	.175	.218	.225	.046	.154	.235	.012	.169	.196	.324***	.299	.385	.446**
Random	F1@5	.160	.169	.175	.072	.103	.138	.012	.159	.165	.208***	.181	.243	.261*
	MRR@10	.261	.300	.314	.109	.159	.220	.023	.260	.278	.361***	.341	.436	.489**
	MAP@10	.279	.315	.335	.116	.171	.240	.024	.279	.296	.389***	.374	.472	.530**
	nDCG@10	.323	.352	.370	.144	.205	.280	.035	.324	.333	.434***	.388	.507	.562**

Table 3: Recommender accuracy results of our two evaluation scenarios. In Scenario 1, we compare approaches that ignore the current tweet content, while in Scenario 2, we compare algorithms that also incorporate the current tweet. We observe that (i) BLL_I outperforms MP_I and MR_I, (ii) BLL_S outperforms MP_S and MR_S, (iii) BLL_{I,S} outperforms MP, FR and CF, and (iv) BLL_{I,S,C} outperforms SR and TCI. Based on a t-test, the symbols * ($\alpha = .1$), ** ($\alpha = .01$) and * ($\alpha = .001$) indicate statistically significant differences between BLL_{I,S} and CF in Scenario 1, and between BLL_{I,S,C} and TCI in Scenario 2.**

proach can be seen as an integrated combination of MP_I and MR_I based on human memory theory.

(c) **MP_S**. The *Most Popular Social Hashtags* algorithm is the social correspondent to the individual MP_I approach [15]. Thus, MP_S does not rank the hashtags based on the frequency in the hashtag assignments of user u but based on the frequency in the hashtag assignments of user u 's set of followers F_u .

(d) **MR_S**. *Most Recent Social Hashtags* is the time-dependent equivalent to MP_S. MR_S sorts the hashtag assignments of u 's followers F_u by time and recommends the k most recent ones. Our BLL_S algorithm is a cognitive-inspired integration of MP_S and MR_S.

(e) **MP**. The unpersonalized *Most Popular Hashtags* approach returns the same set of hashtags for any user. These hashtags are ranked by their overall frequency [15].

(f) **FR**. *FolkRank* is an adaption of Google's PageRank approach used to rank the entities in folksonomy graphs and has become one of the most successful tag recommender methods. We use the standard FR implementation provided by the University of Kassel⁷ with its suggested default parameters. More specifically, the weight of the preference vector d is set to .7 and the maximum number of iterations l is set to 10 [15].

(g) **CF**. *User-based Collaborative Filtering* is a well-known algorithm used in many variants of modern recommender systems and was adapted by [29] for use in tag-based settings. We apply the same idea for the task of recommending hashtags and thus, first identify the k most similar users (i.e., the nearest neighbors) for current user u by means of the cosine similarity measure and then suggest the hashtags used by these neighbors. For our experiments, we use a neighborhood size k of 20 users (see also [8]).

(h) **SR**. *SimilarityRank* is an unpersonalized hashtag recommendation algorithm, which utilizes the content of the currently proposed tweet t [42]. Similarly to our BLL_{I,S,C} approach, this is achieved using TF-IDF to determine content-based similarity scores between tweets (see Section 4.2). These scores are used to recommend the k hashtags that occur in t 's most similar tweets.

(i) **TCI**. *TemporalCombInt* is one of the most recent approaches for personalized hashtag recommendations and also one of the very few approaches that accounts for the effect of

time on hashtag usage [11] (see also Section 6). TCI builds on a linear combination of SR and CF and incorporates temporal effects by considering the time-dependent relevance of a hashtag with respect to the recommendation date. This is done by categorizing the hashtags into "organizational" and "conversational" hashtags, and modeling the decay of temporal relevance using an exponential function. By fitting this model to our crawled data, we set the two main parameters of the algorithm, η_t and η_h , to .1 and .2, respectively.

5.2 Results and Discussion

In Section 3, we found that time is an important factor for hashtag reuse. Because of this, we assume that our time-dependent and cognitive-inspired approach should provide reasonable results compared to other algorithms. The accuracy estimates for our two evaluation scenarios are shown in Table 3 and Figure 4.

Scenario 1: Hashtag rec. w/o current tweet. In our first evaluation scenario, we validate approaches that predict future hashtags without incorporating the content of the currently proposed tweet. Here, we identify three main results:

(a) **BLL_I > MP_I, MR_I**. When predicting individual hashtag reuse, we compare our BLL_I approach to the frequency-based MP_I and the recency-based MR_I algorithms. The results clearly reflect the importance of the time component since MR_I and BLL_I provide higher prediction accuracy and ranking estimates than MP_I for all evaluation metrics across both datasets. Apart from that, we observe that BLL_I outperforms MR_I, which speaks in favor of the cognitive-inspired combination of hashtag frequency and recency by means of the BLL equation.

(b) **BLL_S > MP_S, MR_S**. Concerning the prediction of social hashtag reuse, we compare our BLL_S approach to the frequency-based MP_S and the recency-based MR_S methods. Similar to the case of individual hashtag reuse, MR_S and our BLL-based method provide higher accuracy estimates than the solely frequency-based one, but interestingly, this time the differences between these methods is much larger. This indicates that the time information is especially important in a social setting. We somehow expected this behavior since typically only the most recent tweets of the followers are shown on a user's Twitter timeline. Again, the combination of hashtag frequency and recency by means of the BLL equation provides the best results.

⁷<http://www.kde.cs.uni-kassel.de/code>

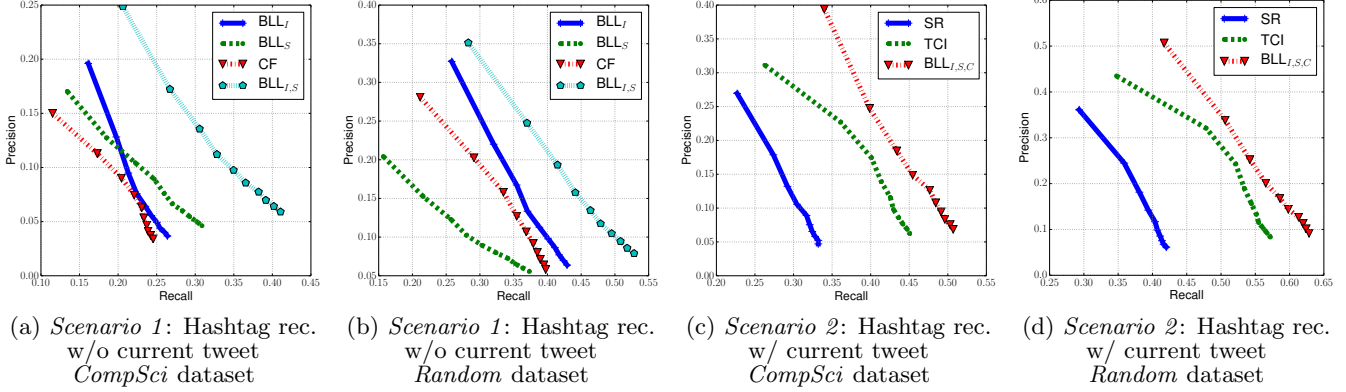


Figure 4: Precision / Recall plots of our two evaluation scenarios showing the accuracy of BLL_I , BLL_S , CF, $BLL_{I,S}$, SR, TCI and $BLL_{I,S,C}$ for $k = 1 - 10$ recommended hashtags. Again, $BLL_{I,S}$ provides the best results in *Scenario 1* and $BLL_{I,S,C}$ provides the best results in *Scenario 2*.

(c) $BLL_{I,S} > MP, FR, CF$. Finally, we compare our hybrid $BLL_{I,S}$ approach to the unpersonalized MP algorithm, the well-known FR method from tag recommender research and classic user-based CF. The first observation that becomes apparent is the poor performance of the unpersonalized MP baseline, which underpins the importance of personalized methods for hashtag recommendation. Additionally, and more importantly, our hybrid $BLL_{I,S}$ approach does not only improve its BLL_I and BLL_S components but also provides significantly higher accuracy and ranking estimates than FR and CF. This shows that $BLL_{I,S}$ is capable of providing reasonable hashtag recommendations solely based on temporal usage patterns of past hashtag assignments.

Scenario 2: Hashtag rec. w/ current tweet. In the second scenario, we evaluate hashtag recommendation methods that also incorporate the content of the current tweet. This includes the unpersonalized SR approach, the time-dependent TCI algorithm and our $BLL_{I,S,C}$ approach. Our two main results are:

(a) $TCI, BLL_{I,S,C} > SR$. The first main result of our second evaluation scenario is that both time-dependent methods TCI and $BLL_{I,S,C}$ outperform the unpersonalized SR approach. We somehow expected this result since both TCI and $BLL_{I,S,C}$ extend the TF-IDF-based tweet content analysis of SR with personalization techniques via CF (TCI) or the BLL equation ($BLL_{I,S,C}$).

(b) $BLL_{I,S,C} > TCI$. The second main result of *Scenario 2* is that $BLL_{I,S,C}$ provides significantly higher accuracy estimates than TCI. This is due to three main differences between these methods: (i) instead of using hashtags of similar users by means of CF for adding personalization, we incorporate not only individual hashtags of the current user but also social hashtags of the current user’s followers, (ii) instead of applying the effect of time on a global hashtag level, we model the time-dependent decay on an individual and social level, and (iii) instead of modeling this time-dependent decay using an exponential function, we use a power function by means of the BLL equation.

CompSci dataset vs. Random dataset. Another interesting finding we observe is that all algorithms provide better results for the *Random* dataset than for the *CompSci* dataset. In our case, this indicates that the task of predict-

ing hashtags in the domain-specific network of computer scientists is harder than in the network of random users. If we look back at Figure 1, this makes sense since the amount of “external” hashtags is twice as high in the *CompSci* dataset (i.e., 26%) than in the *Random* one (i.e., 13%).

Finding 2: The BLL equation, which accounts for temporal effects of item exposure in human memory, provides a suitable model for personalized hashtag recommendations. This is validated in two evaluation scenarios (i.e., without and with incorporating the content of the current tweet), in which our cognitive-inspired approach outperforms several state-of-the-art hashtag recommendation algorithms in terms of prediction accuracy and ranking.

6. RELATED WORK

Over the past years, tagging has emerged as an important feature of the social Web, which supports users to collaboratively organize and find content [17].

Two types of tags have been established: (i) social tags as used in systems like BibSonomy and CiteUlike, and (ii) hashtags as used in systems like Twitter and Instagram. Whereas social tags are mainly used to index resources for later retrieval, hashtags have a more conversational nature and are used to filter and direct content to certain streams of information [12].

Social tag recommendations. One of the most prominent approaches in the field of tag recommendations is the FolkRank algorithm [14, 15]. FolkRank is an extension of the well-known Google PageRank approach to rank the entities in folksonomies (i.e., users, resources and tags). Other important tag recommendation methods are based on Collaborative Filtering [29, 8], Latent Dirichlet Allocation [22, 21] or Tensor Factorization [33, 32]. Recent observations in the field of social tagging state the importance of the time component for the individual tagging behavior of users. In this respect, [43] proposes a time-dependent tag recommender approach, which models the tagging variation over time using an exponential function.

In our previous work [20, 19], we presented a more theory-driven approach, where we use the BLL equation coming from the cognitive architecture ACT-R [3, 2] to model the

power-law of time-dependent decay. We evaluated our approach in detail and compared it to other state-of-the-art methods in [18]. In the present work, we build upon our results and incorporate the BLL equation to study the effect of time on hashtag reuse to design our hashtag recommendation approach.

Hashtag recommendations. There is already a large body of research available that focuses on the recommendation of hashtags in Twitter. One illustrative example is the work presented in [9], in which hashtag recommendations are provided by categorizing tweets into general topics using LDA. The approach then recommends the hashtags that best fit the topics of a new tweet. The authors evaluate their approach using a qualitative study, in which they ask persons if the recommended hashtags describe the topics of a tweet and could be used to semantically enrich it. In 80% of the cases, they are able to provide a suitable hashtag from a selection of five possibilities.

Other similar approaches that use topic models for hashtag recommendations are presented in [37, 40, 7]. In [16], a related algorithm based on a hashtag classification scheme is proposed. The most notable work in the context of hashtag recommendations is probably the content-based SR approach presented in [42]. The authors use the TF-IDF statistic to calculate similarities between tweets and identify suitable hashtags based on these similarity scores. Our $BLL_{I,S,C}$ approach uses the same statistic to integrate the content of a user’s currently proposed tweet. In [24], a personalized extension of SR is presented, in which the authors combine it with user-based CF. Apart from that, a content-based hashtag recommendation algorithm for hyper-linked tweets is proposed in [35].

Temporal effects on hashtag usage. Related research has studied temporal effects on hashtag usage, for instance in the context of popular hashtags in Twitter [26, 25, 39, 28, 27]. For example, in [28], the authors aim to predict if a specific hashtag will be popular on the next day. By formulating this task as a classification problem, they find that both content and context features are effective for popularity prediction. A similar approach is presented in [41], in which the authors uncover the temporal dynamics of online content (e.g., tweets) by formulating a time series clustering problem. One of the very few examples of a time-aware hashtag recommendation approach is the recently proposed algorithm described in [11]. The authors extend the content-based SR approach [42] with a personalization technique by means of CF and further consider the temporal relevance of hashtags. To account for this temporal relevance, they divide the hashtags into two categories: “organizational” ones, which are used over a long period of time and “conversational” ones, which are used only during a short time span (e.g., for a specific event).

In contrast to our proposed algorithm, which relies on the BLL equation, their approach considers the effect of time on a global hashtag level of the whole Twitter network and not on an individual and social level of a specific user. Furthermore, we use a power function rather than an exponential one to model the time-dependent decay.

7. CONCLUSION

In this paper, we presented a cognitive-inspired approach for hashtag recommendations in Twitter. Our approach uti-

lizes the BLL equation from the cognitive architecture ACT-R to account for temporal effects on individual hashtag reuse (i.e., reusing own hashtags) and social hashtag reuse (i.e., reusing hashtags, which has been previously used by a followee). Our analysis of hashtag usage types in two empirical networks (i.e., *CompSci* and *Random* datasets) crawled from Twitter reveals that between 66% and 81% of hashtag assignments can be explained by past individual and social hashtag usage. By analyzing the timestamps of these hashtag assignments, we find that temporal effects play an important role for both individual and social reuse of hashtags and that a power function provides a better fit to model this time-dependent decay than an exponential function.

Thus, the more recently a hashtag was used by a user or her followees, the higher the probability that this user will use the same hashtag again later in time. Based on these findings, we utilized the Base-Level Learning (BLL) equation of the cognitive architecture ACT-R, which accounts for the time-dependent decay of item exposure in human memory, to develop $BLL_{I,S}$ and $BLL_{I,S,C}$, two algorithms for recommending hashtags. Whereas $BLL_{I,S}$ aims to recommend hashtags without incorporating the current tweet (i.e., *Scenario 1*), $BLL_{I,S,C}$ also utilizes the content of the current tweet using the TF-IDF statistic (i.e., *Scenario 2*). We compared both algorithms to state-of-the-art hashtag recommendation algorithms and found that our cognitive-inspired methods outperform these algorithms in terms of prediction accuracy and ranking.

Limitations and future work. One limitation of our work is that we model the reuse of social hashtags solely by analyzing how frequently and recently a hashtag was used by a user’s followees, neglecting by whom the hashtag was used. Thus, for future work, we plan to extend our approach with the social status of the followee (e.g., via the reputation of the user by means of the number of followers). In this respect, we will also utilize the social connection strength between a user and her followee (e.g., by the number of mentions or retweets). Apart from that, it would be interesting to combine our approach with the one presented in [11], thus dividing the hashtags based on their temporal characteristics (i.e., “organizational” vs “conversational”).

Regarding the hashtag assignments that cannot be explained by hashtag reuse (i.e., 26% in the *CompSci* dataset and 13% in the *Random* dataset), we want to utilize an external knowledge base to also account for these hashtag assignments. We will achieve this by suggesting hashtags of currently trending topics or events. Finally, we also plan to verify our findings in larger Twitter data samples than the ones used in this paper as well as in other online social networks that feature the concept of hashtags, such as Instagram and Facebook.

In summary, our work contributes to the rich line of research on improving the use of hashtags in social networks. We hope that future work will be attracted by our insights into how temporal effects on hashtag usage can be modeled using models from human memory theory, such as the BLL equation from the cognitive architecture ACT-R.

Acknowledgments. The authors would like to thank Stefanie Lindstaedt, Tobias Ley, Matthias Traub and Dieter Theiler for valuable inputs. This work is supported by the Know-Center Graz, and the European-funded projects Learning Layers (GA: 318209) and AFEL (GA: 687916).

8. REFERENCES

- [1] J. Alstott, E. Bullmore, and D. Plenz. powerlaw: a python package for analysis of heavy-tailed distributions. *PLoS one*, 9(1):e85777, 2014.
- [2] J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, and Y. Qin. An integrated theory of the mind. *Psychological review*, 111(4):1036, 2004.
- [3] J. R. Anderson and L. J. Schooler. Reflections of the environment in memory. *Psychological Science*, 2(6):396–408, 1991.
- [4] P. G. Campos, F. Díez, and I. Cantador. Time-aware recommender systems: a comprehensive survey and analysis of existing evaluation protocols. *User Modeling and User-Adapted Interaction*, 24(1-2):67–119, 2014.
- [5] A. Clauset, C. R. Shalizi, and M. E. Newman. Power-law distributions in empirical data. *SIAM review (SIREV)*, 51(4):661–703, 2009.
- [6] U. Cress, C. Held, and J. Kimmerle. The collective knowledge of social tags: Direct and indirect influences on navigation, learning, and information processing. *Computers & Education*, 60(1):59–73, 2013.
- [7] M. Efron. Hashtag retrieval in a microblogging environment. In *Proc. of SIGIR'10*, pages 787–788. ACM, 2010.
- [8] J. Gemmel, T. Schimoler, M. Ramezani, L. Christiansen, and B. Mobasher. Improving folkrank with item-based collaborative filtering. *Recommender Systems & the Social Web*, 2009.
- [9] F. Godin, V. Slavkovicj, W. De Neve, B. Schrauwen, and R. Van de Walle. Using topic models for twitter hashtag recommendation. In *Proc. of WWW '13 companion*, pages 593–596. ACM, 2013.
- [10] A. T. Hadgu and R. Jäschke. Identifying and analyzing researchers on twitter. In *Proc. of WebSci '14*, pages 23–30. New York, NY, USA, 2014. ACM.
- [11] M. Harvey and F. Crestani. Long time, no tweets! time-aware personalised hashtag suggestion. In *Proc. of ECIR'15*, pages 581–592. Springer, 2015.
- [12] J. Huang, K. M. Thornton, and E. N. Efthimiadis. Conversational tagging in twitter. In *Proc. of HT '10*. ACM, 2010.
- [13] K. Järvelin and J. Kekäläinen. Ir evaluation methods for retrieving highly relevant documents. In *Proc. of SIGIR'00*, pages 41–48. New York, NY, USA, 2000. ACM.
- [14] R. Jäschke, L. Marinho, A. Hotho, L. Schmidt-Thieme, and G. Stumme. Tag recommendations in folksonomies. In *Proc. of PKDD'07*, pages 506–514. Springer, 2007.
- [15] R. Jäschke, L. Marinho, A. Hotho, L. Schmidt-Thieme, and G. Stumme. Tag recommendations in social bookmarking systems. *AI Communications*, 21(4):231–247, 2008.
- [16] M. Jeon, S. Jun, and E. Hwang. Hashtag recommendation based on user tweet and hashtag classification on twitter. In *Proc. of Web-Age Information Management (WAIM'14)*, pages 325–336. Springer, 2014.
- [17] C. Körner, D. Benz, A. Hotho, M. Strohmaier, and G. Stumme. Stop thinking, start tagging: Tag semantics emerge from collaborative verbosity. In *Proc. of WWW '10*, pages 521–530. New York, NY, USA, 2010. ACM.
- [18] D. Kowald and E. Lex. Evaluating tag recommender algorithms in real-world folksonomies: A comparative study. In *Proc. of RecSys'15*, pages 265–268. New York, NY, USA, 2015. ACM.
- [19] D. Kowald and E. Lex. The influence of frequency, recency and semantic context on the reuse of tags in social tagging systems. In *Proc. of HT'16*, pages 237–242. ACM, 2016.
- [20] D. Kowald, P. Seitlinger, C. Trattner, and T. Ley. Long time no see: The probability of reusing tags as a function of frequency and recency. In *Proc. of WWW'14 companion*, pages 463–468. ACM, 2014.
- [21] R. Krestel and P. Fankhauser. Personalized topic-based tag recommendation. *Neurocomputing*, 76(1):61–70, 2012.
- [22] R. Krestel, P. Fankhauser, and W. Nejdl. Latent dirichlet allocation for tag recommendation. In *Proc. of RecSys'09*, pages 61–68. ACM, 2009.
- [23] H. Kwak, C. Lee, H. Park, and S. Moon. What is twitter, a social network or a news media? In *Proc. of WWW '10*, pages 591–600. New York, NY, USA, 2010. ACM.
- [24] S. M. Kywe, T.-A. Hoang, E.-P. Lim, and F. Zhu. On recommending hashtags in twitter networks. In *Social Informatics*, pages 337–350. Springer, 2012.
- [25] J. Lehmann, B. Gonçalves, J. J. Ramasco, and C. Cattuto. Dynamical classes of collective attention in twitter. In *Proc. of WWW'12*, pages 251–260. ACM, 2012.
- [26] J. Lin and G. Mishne. A study of “churn” in tweets and real-time search queries. In *Proc. of ICWSM'12*, 2012.
- [27] H.-M. Lu and C.-H. Lee. A twitter hashtag recommendation model that accommodates for temporal clustering effects. *IEEE Intelligent Systems*, 30(3):18–25, 2015.
- [28] Z. Ma, A. Sun, and G. Cong. Will this # hashtag be popular tomorrow? In *Proc. of SIGIR'12*, pages 1173–1174. ACM, 2012.
- [29] L. B. Marinho and L. Schmidt-Thieme. Collaborative tag recommendations. In *Data Analysis, Machine Learning and Applications*, pages 533–540. Springer, 2008.
- [30] J. McAuley and J. Leskovec. Hidden factors and hidden topics: understanding rating dimensions with review text. In *Proc. of RecSys'13*, pages 165–172. ACM, 2013.
- [31] S. A. Myers and J. Leskovec. The bursty dynamics of the twitter information network. In *Proc. of WWW'14*, pages 913–924. New York, NY, USA, 2014. ACM.
- [32] S. Rendle, L. Balby Marinho, A. Nanopoulos, and L. Schmidt-Thieme. Learning optimal ranking with tensor factorization for tag recommendation. In *Proc. of KDD'09*, pages 727–736. ACM, 2009.
- [33] S. Rendle and L. Schmidt-Thieme. Pairwise interaction tensor factorization for personalized tag recommendation. In *Proc. of WSDM'10*, pages 81–90. ACM, 2010.
- [34] D. M. Romero, B. Meeder, and J. Kleinberg. Differences in the mechanics of information diffusion across topics: Idioms, political hashtags, and complex contagion on twitter. In *Proc. of WWW '11*, pages 695–704. New York, NY, USA, 2011. ACM.
- [35] S. Sedhai and A. Sun. Hashtag recommendation for hyperlinked tweets. In *Proc. of SIGIR'14*. ACM, 2014.
- [36] P. Seitlinger, T. Ley, and D. Albert. Verbatim and semantic imitation in indexing resources on the web: A fuzzy-trace account of social tagging. *Applied Cognitive Psychology*, 29(1):32–48, 2015.
- [37] J. She and L. Chen. Tomoha: Topic model-based hashtag recommendation on twitter. In *Proc. of WWW '14 companion*, pages 371–372. ACM, 2014.
- [38] T. A. Small. What the hashtag? a content analysis of canadian politics on twitter. *Information, Communication & Society*, 14(6):872–895, 2011.
- [39] O. Tsur and A. Rappoport. What’s in a hashtag?: Content based prediction of the spread of ideas in microblogging communities. In *Proc. of WSDM'12*, pages 643–652. New York, NY, USA, 2012. ACM.
- [40] Y. Wang, J. Qu, J. Liu, J. Chen, and Y. Huang. What to tag your microblog: Hashtag recommendation based on topic analysis and collaborative filtering. In *Web Technologies and Applications*, pages 610–618. Springer, 2014.
- [41] J. Yang and J. Leskovec. Patterns of temporal variation in online media. In *Proc. of WSDM'11*, pages 177–186. ACM, 2011.
- [42] E. Zangerle, W. Gassler, and G. Specht. Recommending#-tags in twitter. In *Proc. of SASWeb'11*, pages 67–78. CEUR Workshop Proceedings, 2011.
- [43] L. Zhang, J. Tang, and M. Zhang. Integrating temporal usage pattern into personalized tag prediction. In *Web Technologies and Applications*. Springer, 2012.