

# Long Time No See: The Probability of Reusing Tags as a Function of Frequency and Recency

Dominik Kowald  
Know-Center  
Graz University of Technology  
Graz, Austria  
dkowald@know-center.at

Christoph Trattner  
Know-Center  
Graz University of Technology  
Graz, Austria  
ctrattner@know-center.at

Paul Seitlinger  
KTI  
Graz University of Technology  
Graz, Austria  
paul.seitlinger@tugraz.at

Tobias Ley  
Institute of Informatics  
Tallinn University  
Tallinn, Estonia  
tley@tlu.ee

## ABSTRACT

In this paper, we introduce a tag recommendation algorithm that mimics the way humans draw on items in their long-term memory. This approach uses the frequency and recency of previous tag assignments to estimate the probability of reusing a particular tag. Using three real-world folksonomies gathered from bookmarks in BibSonomy, CiteULike and Flickr, we show how incorporating a time-dependent component outperforms conventional “most popular tags” approaches and another existing and very effective but less theory-driven, time-dependent recommendation mechanism. By combining our approach with a simple resource-specific frequency analysis, our algorithm outperforms other well-established algorithms, such as FolkRank, Pairwise Interaction Tensor Factorization and Collaborative Filtering. We conclude that our approach provides an accurate and computationally efficient model of a user’s temporal tagging behavior. We demonstrate how effective principles of information retrieval can be designed and implemented if human memory processes are taken into account.

## Categories and Subject Descriptors

H.2.8 [Database Management]: Database Applications—*Data mining*; H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval—*Information filtering*

## Keywords

personalized tag recommendations; time-dependent recommender systems; base-level learning equation; ACT-R; human memory model; BibSonomy; CiteULike; Flickr

## 1. INTRODUCTION

One of the goals of Web Science as a new discipline is to understand the dynamics of human behavior and social interactions that shape the Web into a vast information network of content and people. As the Web evolves into a platform through which people interact with each other, communicate and express themselves, models of human behavior can shed light on why the Web forms as it does and contribute to improve the Web’s underlying mechanisms.

In this paper, we suggest a tag recommendation mechanism that mimics how people use to access their memory to name things they encountered in the past. In everyday communication, people are very effective and quick in retrieving relevant knowledge from the enormous amount of information units stored in their individual long-term memory (LTM). One example is tagging resources on the Web, a rudimentary variant of communication [9, 29]. Here, people name objects, such as images or music files, by means of social tags to create retrieval cues for personal and collective information organization [22]. The issue of how human memory ensures a fast and automatic information retrieval from its huge LTM has been extensively examined by memory psychology (e.g., [1]). Essentially, human memory is tuned to the statistical structure of an individual’s environment and keeps available those memory traces that have been used frequently and recently in the past [2].

Social tagging provides an illustrative example of the strong interplay between external, environmental and internal memory structures (e.g., [11]). For instance, the development of generative models of social tagging demonstrated that the probability of a tag being applied can be modeled through the preferential attachment principle (e.g., [6]): the higher the frequency of a tag’s past occurrence in the tagging environment is, the more likely it will be reused by an individual. Additionally, the same probability is also a function of the tag’s recency, which is the time elapsed since the tag last occurred in the environment [5]. In summary, the probability of applying a particular word reflects the individual’s probability of being exposed to the word in his/her environment [2].

The base-level learning (BLL) equation of the cognitive architecture ACT-R (e.g., [1]) combines the variables of frequency and recency of item exposure to estimate the base-level activation  $BLA_i$

Copyright is held by the International World Wide Web Conference Committee (IW3C2). IW3C2 reserves the right to provide a hyperlink to the author’s site if the Material is used in electronic media.  
*WWW’14 Companion*, April 7–11, 2014, Seoul, Korea.  
ACM 978-1-4503-2745-9/14/04.  
<http://dx.doi.org/10.1145/2567948.2576934>.

of the memory trace  $i$  for the corresponding item. It is given by

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

, where  $n$  is the frequency of item occurrences in the past and  $t_j$  is the recency, which is the time since the  $j^{th}$  occurrence. The exponent  $d$  accounts for the power-law of forgetting and models the phenomenon that each memory’s activation, caused by the  $j^{th}$  occurrence, decreases in time according to a power function. The exponent  $d$  is typically set to 0.5 [1].

Based on the research briefly described above, we assume that a user’s past usage of a tag predicts the probability of its future usage. Hence, equation (1) should help to infer the probability of a tag being applied during a new tag assignment. In particular, if frequency and recency are both strong predictors of a tag’s reuse probability, the base-level learning equation should help to extend simple “most popular tags” approaches based solely on frequency analyses.

The work of [33] provides empirical proof of this assumption. They showed that a recommender, which combines the frequency and recency of tag use, has a higher accuracy with respect to recall and precision than a recommender only taking into account the frequency of tag use. However, the equations they used to implement their approach were developed from scratch and not derived from existing research described above (see Sections 4.3 and 6).

The research questions of this work are as follows: (i) Does the BLL equation provide a valid model of a user’s tagging behavior in the past to predict future tag assignments? (ii) Can the BLL equation, that integrates frequency and recency of tag usage, be applied and extended to create an effective and efficient tag recommendation mechanism?

The strategy we chose to address both research questions consisted of two steps. In a first step, we implemented the “pure” BLL equation in form of a tag recommender and compared its performance with a MostPopular<sub>*u*</sub> (MP<sub>*u*</sub>) approach suggesting the most frequent tags in a user’s tag assignments. This comparison should reveal the increment value that may result from additionally processing the recency of tag use. Moreover, we compared our BLL recommender with the approach introduced by [33] in order to determine potential advantages of our theory-driven approach.

In a second step, we extended the BLL equation to also take into account the effect of popular tags (i.e., semantic cues  $C$ ) associated with a resource on the availability of memory traces and hence, tagging behavior. For the first approximation of  $C$  we decided to simply weight the tags based on their frequency in the tag assignments of the resource (hereinafter referred to as MostPopular<sub>*r*</sub> (MP<sub>*r*</sub>)). We then compared the performance of this combination of BLL and MP<sub>*r*</sub> (BLL+C) with well-established approaches, such as Collaborative Filtering (CF), FolkRank (FR) and Pairwise Interaction Tensor Factorization (PITF), to examine our second research question.

The remainder of this paper is organized as follows. We begin with discussing related work (Section 2) and describing our approach in Section 3. Sections 4 and 5 address our two research questions and summarize the settings and results of our extensive evaluation. Finally, in Section 6, we conclude the paper by discussing our findings in the light of the benefits of deriving tag recommender mechanisms from empirical, cognitive research.

## 2. RELATED WORK

Recent years have shown that tagging is an important feature of the Social Web supporting the users with a simple mechanism to

collaboratively organize and find content [16]. Although tagging has the ability to improve search (and in particular tags provided by the individual), it is also known that users are typically lazy in providing tags for instance for their bookmarked resources. It is therefore not surprising that recent research has attempted to address this challenge to support the individual in her tag application process in the form of personalized tag recommenders. To date, the two following approaches have been established: graph based and content-based tag recommender systems [18]. In our work we focus on graph-based approaches.

The probably most notable work in this context is the work of Hotho et al. [13] who introduced an algorithm termed FolkRank (FR) that has become the most prominent benchmarking tag recommender approach over the past few years. Subsequently, the work of Jäschke et al. [14] and Hamouda & Wanas [10] showed how the classic Collaborative Filtering (CF) approach could be adopted for the problem of predicting tags to the user in a personalized manner. More recent work in this context are studies of Rendle et al. [26], Wetzker et al. [30], Krestel et al. [17] or Rawashdeh et al. [24] who introduced a factorization model, a Latent Dirichlet Allocation (LDA) model or a Link-Prediction model, based on the Katz measure, to recommend tags to users.

Although the latter mentioned approaches perform reasonably well, they are computationally expensive compared to simple “most popular tags” approaches. Furthermore, they ignore recent observations with regard to social tagging systems, such as the variation of the individual tagging behavior over time [32]. To that end, recent research has made the first promising steps towards more accurate graph-based models that also account for the variable of time [31, 33]. The approaches have shown to outperform some of the current state-of-the-art tag recommender algorithms.

In line with the latter strand of research, in this paper we present a novel graph-based tag recommender mechanism that uses the BLL equation which is based on the principles of a popular model of human cognition called ACT-R (e.g., [1]). We show that the approach is not only very simple and straightforward but also reveal that the algorithm outperforms current state-of-the-art graph-based (e.g., [30, 13, 14]) and the leading time-based [33] tag recommender approaches.

## 3. APPROACH

In Section 1 we formulated the assumption that both frequency and recency of tag use explain a high variance in a tag’s probability of being applied and that this probability could be modeled via the BLL equation introduced by Anderson et al. [1]. In the following we describe how we implemented the BLL equation to calculate the base-level activation (BLA) of a given tag  $t$  in a user’s set of tag assignments,  $Y_{t,u}$ . First, we determined a reference timestamp  $timestamp_{ref}$  (in seconds) that is the timestamp of the most recent bookmark of user  $u$ . In our dataset samples,  $timestamp_{ref}$  corresponded to the timestamp of  $u$ ’s bookmark in the test set (see Section 4.1).

If  $i = 1 \dots n$  indexes all tag assignments in  $Y_{t,u}$ , the recency of a particular tag assignment is given by  $timestamp_{ref} - timestamp_i$ . Finally, the BLA of tag  $t$  for a user  $u$  is given by the BLL equation:

$$BLA(t, u) = \ln\left(\sum_{i=1}^n (timestamp_{ref} - timestamp_i)^{-d}\right) \quad (2)$$

, where  $d$  is set to 0.5 based on [1]. In order to map the values onto a range of 0 - 1 we applied a normalization method as proposed in

Dataset	Core	$ B $	$ U $	$ R $	$ T $	$ TAS $
BibSonomy	-	400,983	5,488	346,444	103,503	1,479,970
	3	41,764	788	8,711	5,757	161,509
CiteULike	-	3,879,371	83,225	2,955,132	800,052	16,703,839
	3	735,292	17,983	149,220	67,072	2,242,849
Flickr	-	864,679	9,590	864,679	127,599	3,552,540
	3	860,135	8,332	860,135	58,831	3,465,346

**Table 1: Properties of the datasets, where  $|B|$  is the number of bookmarks,  $|U|$  the number of users,  $|R|$  the number of resources,  $|T|$  the number of tags and  $|TAS|$  the number of tag assignments.**

related work [23]:

$$\|BLA(t, u)\| = \frac{\exp(BLA(t, u))}{\sum_{t'=1}^m \exp(BLA(t', u))} \quad (3)$$

, where  $m$  equals  $|Y_u|$ .

When incorporating BLL into a recommender, we aim at predicting the probability of a word being applied in the present tag assignment. To this end, we also have to take into account semantic cues  $C$  in a user’s current environment (e.g., the resource to be tagged) to fine-tune the “prior” probability estimated via the BLL equation (e.g., [1]). In the case of tagging a resource,  $C$  partially consists of content words in the title and in the page text or of prominent tags associated with the resource (e.g., [19, 18]). Since we focus in this work on graph-based approaches and not all of our datasets contain title information or page-text, we modeled the influence of  $C$  by simply taking into account the most popular tags of the resource ( $MP_r$ , i.e.,  $\arg \max_{t \in T} (|Y_{t,r}|)$ ), which is also normalized using equation (3) [13]. Thus, we applied  $MP_r$  to adjust the BLA of a given tag according to potential semantic cues available in the user’s environment. Taken together, the list of recommended tags for a given user  $u$  and resource  $r$  is calculated by

$$\tilde{T}(u, r) = \arg \max_{t \in T} \underbrace{(\beta \|BLA(t, u)\|)}_{BLL} + \underbrace{(1 - \beta) \|Y_{t,r}\|}_{BLL+C} \quad (4)$$

, where  $\beta$  is used to inversely weight the two components, i.e. the BLA and the semantic cues  $C$ . The results presented in Section 5 were calculated using  $\beta = 0.5$ . However, we focused on the performance of BLL+C in the experiments, i.e. on an approach estimating a tag’s probability of being applied by means of user and resource information. Taken together, this is in line with the ACT-R declarative module that also considers retrieval probability as a function of base-level activation and environmental features.

## 4. EXPERIMENTAL SETUP

In this section we describe in detail the datasets, the evaluation method, the metrics and the algorithms used in our experiments.

### 4.1 Datasets

For the purposes of our study and for reasons of reproducibility, we focused our investigations on three well-known and freely-available folksonomy datasets. To test our approach on both types of advocates, known as broad and narrow folksonomies [12] (in a broad folksonomy many users are allowed to annotate a particular resource while in a narrow folksonomy only the user who has uploaded the resource is permitted to apply tags), freely available datasets from the social bookmark and publication sharing system

Dataset	Core	Measure	$MP_u$	GIRP	BLL
BibSonomy	-	$F_1@5$	.152	.157	<b>.162</b>
		MRR	.114	.119	<b>.125</b>
		MAP	.148	.155	<b>.162</b>
	3	$F_1@5$	.215	.221	<b>.228</b>
		MRR	.202	.210	<b>.230</b>
		MAP	.238	.247	<b>.272</b>
CiteULike	-	$F_1@5$	.185	.194	<b>.201</b>
		MRR	.165	.182	<b>.193</b>
		MAP	.194	.213	<b>.227</b>
	3	$F_1@5$	.272	.291	<b>.300</b>
		MRR	.268	.294	<b>.319</b>
		MAP	.305	.337	<b>.366</b>
Flickr	-	$F_1@5$	.435	.509	<b>.523</b>
		MRR	.360	.445	<b>.466</b>
		MAP	.468	.590	<b>.619</b>
	3	$F_1@5$	.488	.577	<b>.592</b>
		MRR	.407	.511	<b>.533</b>
		MAP	.527	.676	<b>.707</b>

**Table 2:  $F_1@5$ , MRR and MAP values for BibSonomy, CiteULike and Flickr (no core and core 3) showing that our BLL equation provides a valid model of a user’s tagging behavior to predict tags (first research question).**

BibSonomy<sup>1</sup>, the reference management system CiteULike<sup>2</sup> (broad folksonomies) and the image and video sharing platform Flickr<sup>3</sup> (narrow folksonomy) were utilized. Since automatically generated tags affect the performance of the tag recommender systems, we excluded all of those tags from the datasets (e.g., for BibSonomy and CiteULike we excluded the *no-tag*, *bibtex-import-tag*, etc.). Furthermore, we decapitalized all tags as suggested by related work in the field (e.g., [26]). In the case of Flickr we randomly selected 3% of the user profiles for reasons of computational effort (see also [8]). The overall dataset statistics can be found in Table 1. As depicted, we applied both: a  $p$ -core pruning approach [3] to capture the issues of data sparseness, as well as no  $p$ -core pruning to capture the issue of cold-start users or items [7], respectively.

### 4.2 Evaluation Methodology

To evaluate our tag recommender approach we used a leave-one-out hold-out method as proposed by popular and related work in this area (e.g., [14]). Hence, we created two sets, one for training and the other for testing. To split up each dataset in those two sets we eliminated for each user his/her latest bookmark (in time) from the original dataset and added it to the test set. The remaining original dataset was then used for training, and the newly created one for testing. This procedure simulates a real-world environment well and is a recommended offline-evaluation procedure for time-based recommender systems [4]. To finally quantify the performance of our approach, a set of well-known information retrieval performance standard metrics was used. In particular, we report in Section 5 Recall ( $R@k$ ), Precision ( $P@k$ ), F1-Score ( $F_1@k$ ), Mean Reciprocal Rank (MRR) and Mean Average Precision (MAP), where  $k$  is between 1 and 10 and MRR and MAP are calculated for 10 recommended tags ( $k = 10$ ) [14, 18].

### 4.3 Baseline Algorithms

We compared the results of our approach to several baseline tag recommender algorithms. The algorithms were selected based on their popularity in the community, performance and novelty [20]:

<sup>1</sup><http://www.kde.cs.uni-kassel.de/bibsonomy/dumps>

<sup>2</sup><http://www.citeulike.org/faq/data.asp>

<sup>3</sup><http://www.tagora-project.eu/>

Dataset	Core	Measure	MP	MP <sub>r</sub>	MP <sub>u,r</sub>	CF	APR	FR	FM	PITF	GIRPTM	BLL+C
BibSonomy	-	F <sub>1</sub> @5	.013	.074	.192	.166	.175	.171	.122	.139	.197	<b>.201</b>
		MRR	.008	.054	.148	.133	.149	.148	.097	.120	.152	<b>.158</b>
		MAP	.009	.070	.194	.173	.193	.194	.120	.150	.200	<b>.207</b>
	3	F <sub>1</sub> @5	.047	.313	.335	.325	.260	.337	.345	.356	.350	<b>.353</b>
		MRR	.035	.283	.327	.289	.279	.333	.329	.341	.334	<b>.349</b>
		MAP	.038	.345	.403	.356	.329	.414	.408	.421	.416	<b>.435</b>
CiteULike	-	F <sub>1</sub> @5	.002	.131	.253	.218	.195	.194	.111	.122	.263	<b>.270</b>
		MRR	.001	.104	.229	.201	.233	.233	.110	.141	.246	<b>.258</b>
		MAP	.001	.134	.280	.247	.284	.284	.125	.158	.301	<b>.315</b>
	3	F <sub>1</sub> @5	.013	.270	.316	.332	.313	.318	.254	.258	.336	<b>.346</b>
		MRR	.012	.243	.353	.295	.361	.366	.282	.290	.380	<b>.409</b>
		MAP	.012	.294	.420	.363	.429	.436	.326	.334	.455	<b>.489</b>
Flickr	-	F <sub>1</sub> @5	.023	-	.435	.417	.328	.334	.297	.316	.509	<b>.523</b>
		MRR	.023	-	.360	.436	.352	.355	.300	.333	.445	<b>.466</b>
		MAP	.023	-	.468	.581	.453	.459	.384	.426	.590	<b>.619</b>
	3	F <sub>1</sub> @5	.026	-	.488	.493	.368	.378	.361	.369	.577	<b>.592</b>
		MRR	.026	-	.407	.498	.398	.404	.375	.390	.511	<b>.533</b>
		MAP	.026	-	.527	.663	.513	.523	.481	.502	.676	<b>.707</b>

**Table 3:**  $F_1@5$ , MRR and MAP values for BibSonomy, CiteULike and Flickr (no core and core 3) showing that BLL+C outperforms state-of-the-art baseline algorithms (second research question).

**MostPopular (MP):** This approach recommends for any user and any resource the same set of tags that is weighted by the frequency in all tag assignments [15].

**MostPopular<sub>u</sub> (MP<sub>u</sub>):** The *most popular tags by user* approach suggests the most frequent tags in the tag assignments of the user [15].

**MostPopular<sub>r</sub> (MP<sub>r</sub>):** The *most popular tags by resource* algorithm weights the tags based on their frequency in the tag assignments of the resource [15].

**MostPopular<sub>u,r</sub> (MP<sub>u,r</sub>):** This algorithm is a mixture of the most popular tags by user (MP<sub>u</sub>) and most popular tags by resource (MP<sub>r</sub>) approaches [14].

**Collaborative Filtering (CF<sub>u</sub>):** Marinho et al. [21] described how the classic Collaborative Filtering (CF) approach [27] can be adapted for tag recommendations, where the neighborhood of an user is formed based on the tag assignments in the user profile. The only variable parameter here is the number of users in the neighborhood which has been set to 20 according to the work of Gemmill et al. [8].

**Adapted PageRank (APR):** Hotho et al. [13] adapted the well-known PageRank algorithm in order to rank the nodes within the graph structure of a folksonomy. This is based on the idea that a resource is important if it is tagged with important tags by important users.

**FolkRank (FR):** The FolkRank algorithm is an extension of the Adapted PageRank approach that gives a higher importance to the preference vector via a differential approach [14]. Our APR and FR implementations are based on the code and settings of the open-source Java tag recommender framework provided by the University of Kassel<sup>4</sup>.

**Factorization Machines (FM):** Rendle [25] introduced Factorization Machines which combine the advantages of Support Vector Machines (SVM) with factorization machines to build a general prediction model that is also capable of tag recommendations.

**Pairwise Interaction Tensor Factorization (PITF):** This approach proposed by Rendle and Schmidt-Thieme [26] is an extension of factorization models based on the Tucker Decomposition (TD) model that explicitly models the pairwise interactions between users, resources and tags. The FM and PITF results presented in this paper were calculated using the open-source C++ tag

recommender framework provided by the University of Konstanz<sup>5</sup> with 256 factors, as suggested by [26].

**Temporal Tag Usage Patterns (GIRP):** This time-dependent tag-recommender algorithm was presented by Zhang et al. [33] and is based on the frequency and the temporal usage of a user’s tag assignments. In contrast to BLL it models the temporal tag usage with an exponential distribution rather than a power-law distribution.

**GIRP with Tag Relevance to Resource (GIRPTM):** This is an extension of the GIRP algorithm that also takes the resource component into account in the same manner as in BLL+C [33].

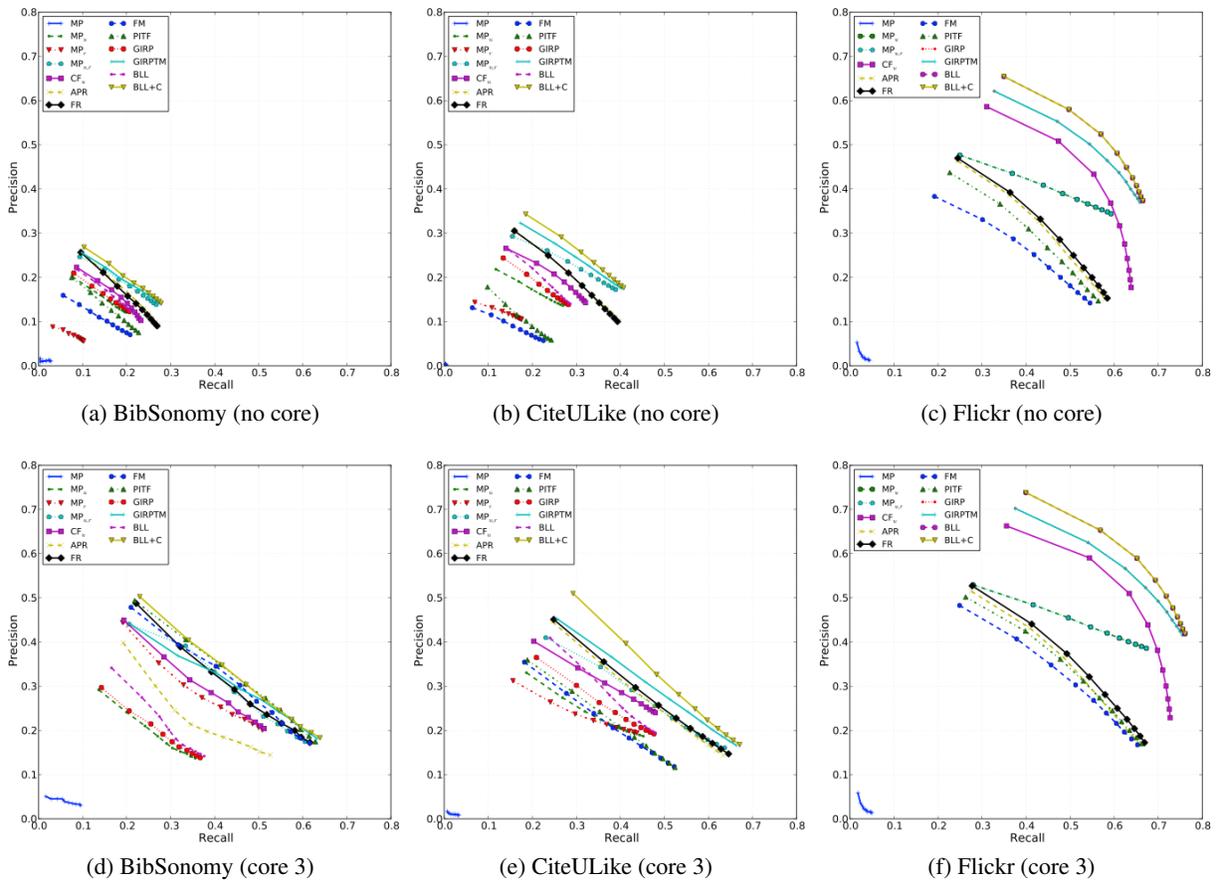
## 5. RESULTS

In this section we present the results of our experiments with respect to the recommender quality in two steps as described in Section 1. In the first step we compared BLL with MP<sub>u</sub> and GIRP in order to determine the effect of the recency component of the tag assignments. The results in Table 2 clearly show that both time-dependent algorithms outperform the frequency-based MP<sub>u</sub> approach. Furthermore, BLL reaches higher levels of accuracy than the less theory-driven GIRP algorithm in both settings (without using a core and with core 3). Moreover, it becomes apparent that the impact of the recency component is significantly higher in the narrow folksonomy (Flickr), where BLL and BLL+C are equal, than in the broad folksonomies (BibSonomy, CiteULike).

In the second step we compared our BLL+C approach, which is a combination of BLL and MP<sub>r</sub>, with a set of state-of-the-art baseline algorithms. When looking at the results in Table 3, the first thing that comes apparent is the fact that the two time-dependent algorithms (GIRPTM and BLL+C) produce the highest estimates ( $F_1@5$ , MRR and MAP) across all three datasets and in both settings (with  $p$ -core pruning applied and without). Second, all recommender algorithms substantially outperform the baseline mechanism, i.e., the simplest “most popular tags” approach MP. Third, our BLL+C approach also outperforms GIRPTM, the currently leading graph-based time-dependent tag recommendation algorithm, especially in terms of the ranking-dependent metrics, such as MRR and MAP. Same observations can be made when looking at the Recall / Precision curves in Figure 1.

<sup>4</sup><http://www.kde.cs.uni-kassel.de/code>

<sup>5</sup><http://www.informatik.uni-konstanz.de/rendle/software/tag-recommender/>



**Figure 1: Recall/Precision plots for BibSonomy, CiteULike and Flickr (no core and core 3) showing the performance of the algorithms for 1 - 10 recommended tags ( $k$ ).**

In summary, this pattern of results implies that the base-level learning equation can be used to implement a highly effective recommender approach. By considering the recency in addition to the frequency of tag use with the help of this equation as well as the current context, it exceeds the performance of well-established and effective recommenders, such as  $MP_{u,r}$ , CF, APR, FM and the other time-dependent approach GIRPTM. Most surprisingly, despite its simplicity, BLL+C appears to be even more successful than the sophisticated FR and PITF algorithms. The code we used for our experiments is open-source and can be found online<sup>6</sup>.

## 6. DISCUSSION AND CONCLUSION

In this study we have followed a two-step strategy and started by comparing the performance of BLL with  $MP_u$  to determine the effect of additionally considering the recency of each tag use as well as with GIRP to differentiate our cognitive-psychological model with the less theory-driven approach introduced by Zhang et al. [33] in order to tackle our first research question. Our results clearly demonstrate that, regardless of the evaluation metric and across all datasets, BLL reaches higher levels of accuracy than  $MP_u$  and outperforms GIRP. Thus, processing the recency of tag use is effective to account for additional variance of users' tagging behavior and therefore, a reasonable extension of simple "most popular tags" approaches. Furthermore, the advantage over GIRP indicates that

drawing on memory psychology guides the application of a reliable and valid model built upon long-standing, empirical research. The equations that Zhang et al. [33] used to implement their approach were developed from scratch rather than derived from existing research described above. As a consequence, [33] modeled the recency of tag use by means of an exponential function, which is clearly at odds with the power law of forgetting described in the introduction.

In a second step, we have combined BLL with  $MP_r$ . Where BLL gives the prior probability of tag reuse that is learned over time,  $MP_r$  tunes this prior probability to the current context by exploiting the current semantic cues from the environment. This is in line with how ACT-R models the retrieval from long-term memory. Despite its simplicity, our results show that this combination (BLL+C) has potential to outperform well-established mechanisms, such as CF, FR and PITF. We assume this is the case because, in following some fundamental principles of human memory, BLL+C is better adapted to the statistical structure of the environment.

Moreover, a glance on the results shows an interaction between the examined dataset and the performance of BLL (and BLL+C). While the distance to other strongly performing mechanisms does not appear to be large in case of broad folksonomies (BibSonomy and CiteULike), this distance gets substantially larger in a narrow folksonomy (Flickr), where BLL and BLL+C have the same performance. From this interaction we conclude that applying a model of human memory is primarily effective if tag assignments

<sup>6</sup><https://github.com/domkowald/tagrecommender>

are mainly driven by individual habits unaffected by the behavior of other users, such as it is done in Flickr.

In future work, we will continue examining memory processes involved in categorizing and tagging Web resources. For instance, in a recent study [28], we introduced a mechanism by which memory processes involved in tagging can be modeled on two levels of knowledge representation: on a semantic level (representing categories or LDA topics) and on a verbal level (representing tags). Next, we will aim at combining this integrative mechanism with the BLL equation to examine a potential interaction between the impact of recency (time-based forgetting) and the level of knowledge representation. As before, conclusions drawn from cognitive science should help to develop an effective and psychologically plausible tag recommendation mechanism.

**Acknowledgments:** This work is supported by the Know-Center, the EU funded project Learning Layers (Grant Agreement 318209) and the Austrian Science Fund (FWF): P 25593-G22.

## 7. REFERENCES

- [1] J. R. Anderson, M. D. Byrne, S. Douglass, C. Lebiere, and Y. Qin. An integrated theory of the mind. *Psychological Review*, 111(4):1036–1050, 2004.
- [2] J. R. Anderson and L. J. Schooler. Reflections of the environment in memory. *Psychological Science*, 2(6):396–408, 1991.
- [3] V. Batagelj and M. Zaveršnik. Generalized cores. *arXiv preprint cs/0202039*, 2002.
- [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*, pages 1–53, 2013.
- [5] C. Cattuto, V. Loreto, and L. Pietronero. Semiotic dynamics and collaborative tagging. *Proceedings of the National Academy of Sciences*, 104(5):1461–1464, 2007.
- [6] K. Dellschaft and S. Staab. An epistemic dynamic model for tagging systems. In *Proceedings of the nineteenth ACM conference on Hypertext and hypermedia*, HT '08, pages 71–80, New York, NY, USA, 2008. ACM.
- [7] S. Doerfel and R. Jäschke. An analysis of tag-recommender evaluation procedures. In *Proceedings of the 7th ACM Conference on Recommender Systems*, RecSys '13, pages 343–346, New York, NY, USA, 2013. ACM.
- [8] J. Gemmell, T. Schimoler, M. Ramezani, L. Christiansen, and B. Mobasher. Improving folkrank with item-based collaborative filtering. *Recommender Systems & the Social Web*, 2009.
- [9] H. Halpin, V. Robu, and H. Shepherd. The complex dynamics of collaborative tagging. In *Proceedings of the 16th international conference on World Wide Web*, WWW '07, pages 211–220, New York, NY, USA, 2007. ACM.
- [10] S. Hamouda and N. Wanas. Put-tag: personalized user-centric tag recommendation for social bookmarking systems. *Social network analysis and mining*, 1(4):377–385, 2011.
- [11] C. Held, J. Kimmerle, and U. Cress. Learning by foraging: The impact of individual knowledge and social tags on web navigation processes. *Computers in Human Behavior*, 28(1):34 – 40, 2012.
- [12] D. Helic, C. Körner, M. Granitzer, M. Strohmaier, and C. Trattner. Navigational efficiency of broad vs. narrow folksonomies. In *Proceedings of the 23rd ACM conference on Hypertext and social media*, HT '12, pages 63–72, New York, NY, USA, 2012. ACM.
- [13] A. Hotho, R. Jäschke, C. Schmitz, and G. Stumme. Information retrieval in folksonomies: Search and ranking. In *The semantic web: research and applications*, pages 411–426. Springer, 2006.
- [14] R. Jäschke, L. Marinho, A. Hotho, L. Schmidt-Thieme, and G. Stumme. Tag recommendations in folksonomies. In *Knowledge Discovery in Databases: PKDD 2007*, 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] C. Körner, D. Benz, A. Hotho, M. Strohmaier, and G. Stumme. Stop thinking, start tagging: tag semantics emerge from collaborative verbosity. In *Proceedings of the 19th international conference on World wide web*, WWW '10, pages 521–530, New York, NY, USA, 2010. ACM.
- [17] R. Krestel and P. Fankhauser. Language models and topic models for personalizing tag recommendation. In *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/ACM International Conference on*, volume 1, pages 82–89. IEEE, 2010.
- [18] M. Lipczak. *Hybrid Tag Recommendation in Collaborative Tagging Systems*. PhD thesis, Dalhousie University, 2012.
- [19] J. Lorince and P. M. Todd. Can simple social copying heuristics explain tag popularity in a collaborative tagging system? In *Proceedings of the 5th Annual ACM Web Science Conference*, WebSci '13, pages 215–224, New York, NY, USA, 2013. ACM.
- [20] L. Marinho, A. Nanopoulos, L. Schmidt-Thieme, R. Jäschke, A. Hotho, G. Stumme, and P. Symeonidis. Social tagging recommender systems. In F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, editors, *Recommender Systems Handbook*, pages 615–644. Springer US, 2011.
- [21] L. B. Marinho and L. Schmidt-Thieme. Collaborative tag recommendations. In *Data Analysis, Machine Learning and Applications*, pages 533–540. Springer, 2008.
- [22] C. Marlow, M. Naaman, D. Boyd, and M. Davis. Ht06, tagging paper, taxonomy, flickr, academic article, to read. In *Proceedings of the seventeenth conference on Hypertext and hypermedia*, HYPERTEXT '06, pages 31–40, New York, NY, USA, 2006. ACM.
- [23] J. McAuley and J. Leskovec. Hidden factors and hidden topics: Understanding rating dimensions with review text. In *Proceedings of the ACM Conference Series on Recommender Systems*, New York, NY, USA, 2013. ACM.
- [24] M. Rawashdeh, H.-N. Kim, J. M. Alja'am, and A. El Saddik. Folksonomy link prediction based on a tripartite graph for tag recommendation. *Journal of Intelligent Information Systems*, pages 1–19, 2012.
- [25] S. Rendle. Factorization machines. In *Data Mining (ICDM), 2010 IEEE 10th International Conference on*, pages 995–1000. IEEE, 2010.
- [26] S. Rendle and L. Schmidt-Thieme. Pairwise interaction tensor factorization for personalized tag recommendation. In *Proceedings of the third ACM international conference on Web search and data mining*, WSDM '10, pages 81–90, New York, NY, USA, 2010. ACM.
- [27] J. B. Schafer, D. Frankowski, J. Herlocker, and S. Sen. Collaborative filtering recommender systems. In *The adaptive web*, pages 291–324. Springer, 2007.
- [28] P. Seitlinger, D. Kowald, C. Trattner, and T. Ley. Recommending tags with a model of human categorization. In *The ACM International Conference on Information and Knowledge Management*, New York, NY, USA, 2013. ACM.
- [29] L. Steels. Semiotic dynamics for embodied agents. *IEEE Intelligent Systems*, 21(3):32–38, 2006.
- [30] R. Wetzker, C. Zimmermann, C. Bauckhage, and S. Albayrak. I tag, you tag: translating tags for advanced user models. In *Proceedings of the third ACM international conference on Web search and data mining*, pages 71–80. ACM, 2010.
- [31] D. Yin, L. Hong, and B. D. Davison. Exploiting session-like behaviors in tag prediction. In *Proceedings of the 20th international conference companion on World wide web*, pages 167–168. ACM, 2011.
- [32] D. Yin, L. Hong, Z. Xue, and B. D. Davison. Temporal dynamics of user interests in tagging systems. In *Twenty-Fifth AAAI conference on artificial intelligence*, 2011.
- [33] L. Zhang, J. Tang, and M. Zhang. Integrating temporal usage pattern into personalized tag prediction. In *Web Technologies and Applications*, pages 354–365. Springer, 2012.