FolkTrails: Interpreting Navigation Behavior in a Social Tagging System

Thomas Niebler, Martin Becker, Daniel Zoller
DMIR Group, University of Würzburg, Germany
{niebler,becker,zoller}@informatik.uni-wuerzburg.de

Stephan Doerfel
KDE Group, University of Kassel, Germany
doerfel@cs.uni-kassel.de

Andreas Hotho
DMIR Group, University of Würzburg and L3S Research Center, Germany
hotho@informatik.uni-wuerzburg.de

ABSTRACT

Social tagging systems have established themselves as a quick and easy way to organize information by annotating resources with tags. In recent work, user behavior in social tagging systems was studied, that is, how users assign tags, and consume content. However, it is still unclear how users make use of the navigation options they are given. Understanding their behavior and differences in behavior of different user groups is an important step towards assessing the effectiveness of a navigational concept and improving it to better suit the users' needs. In this work, we investigate navigation trails in the popular scholarly social tagging system BibSonomy from six years of log data. We discuss dynamic browsing behavior of the general user population and show that different navigational subgroups exhibit different navigational traits. Furthermore, we provide strong evidence that the semantic nature of the underlying folksonomy is an essential factor for explaining navigation.

1. INTRODUCTION

With the advent of social media and the Web 2.0, it became very easy to contribute and arrange content on web platforms, even for technical laymen, for example in wikis, social networks or social tagging systems. As a direct consequence, not only providing, but browsing and consuming content accounts for a big part of user interaction in today's Web. Among those various platforms, social tagging systems have established themselves as a quick and easy way to organize and store information, such as websites[1] and publications[2]. In such systems, users can post resources and freely annotate them with keywords (called *tags*), for example, for means of later retrieval by themselves and by other users. The emerging structure over users, tags, and resources and their connections is called a *folksonomy* and serves as the main navigational concept in social tagging systems, providing links between co-occurring entities. Through those links, folksonomies possess an inherently semantic nature. However, it is still largely unclear how users make use of the navigation options they are given. Understanding their behavior and its differences between various user groups is an important step towards assessing the effectiveness of a navigational concept and of improving it to better suit the users' needs. This task has attracted broad interest in the research community and previous work has focused on the navigation within one particular website [20, 7] or on the Web in general [6, 14]. Also, studying the navigation behavior of users in social tagging systems is of great interest, especially because of these systems' inherent semantic structure. Consequently, different studies have addressed this issue: [9] conducted a user survey on usage motivation in social tagging systems and [7] used log files to study actual navigation behavior of the overall user population through request counts. While these findings give a first insight into the behavioral properties of user navigation, they focus on the overall population of users and their *global* request counts.

Problem Setting and Approach: Although the above mentioned works give first insights into behavioral properties of user navigation, there exist several competing ideas and hypotheses about how users browse within a social tagging system based on *local* transition probabilities. Up to this point, such hypotheses were not objectively compared on actual navigation data. To address this issue, we utilize log files of BibSonomy, which provide a unique opportunity to study the navigational trails of user groups in a folksonomy. In particular, we formulate several navigation hypotheses and compare them using HypTrails [17], a method for comparing hypotheses about human movement on the Web. We also study how the performance of explaining navigation behavior differs on different data subsets, such as navigation grouped by gender, tagging behavior, or long-term experience. In the process, we revisit the aspects described in [7], and extend on their work, providing additional explanations for user intentions during navigation and their comparison. Furthermore, because tags and their inherent semantic information exert a great influence on social tagging systems, it is a logical assumption that navigating those systems is influenced by the semantic content. Thus, we explicitly search for a signal of the influence of tags on navigation, that is, a semantic component.

Contributions: The contributions in this paper are threefold: 1. We study different hypotheses about the navigational user behavior in a tagging system. 2. We provide evidence for the existence of a semantic influence on navigation. 3. We show that users with different tagging behavior also exhibit different navigational traits.

Overall, we contribute to a better understanding of navigation in tagging systems and the role of the folksonomy structure. We support the claims from [7] and shed light on general as well as subgroup specific behavior. We expect our results to be relevant not only for researchers interested in understanding human behavior and social tagging, or operators of any system utilizing tags (e.g., Twitter), but also for the whole Semantic Web community.

[1] e.g., Delicious (http://www.delicious.com)

[2] e.g., BibSonomy (http://www.bibsonomy.org)

CIKM'16 , October 24 - 28, 2016, Indianapolis, IN, USA

ISBN 978-1-4503-4073-1/16/10…$15.00

DOI: http://dx.doi.org/10.1145/2983323.2983686

2. RELATED WORK

In the following, we cover related work on research about the analysis of user navigation in web systems, as well as, research covering folksonomies and social tagging systems.

Navigation Analysis: Understanding human trails on the Web and the underlying generative processes is an open and complex challenge for the research community. However, a lot of previous work addresses these problems. In [20], click paths from a navigation game on Wikipedia have been examined to find way-finding patterns. A promising approach to model human navigation is provided by Markov models, which have been previously used for click-stream data in [4]. [18] noted that while Markov chains of higher order are too complex and inefficient to actually be useful, the memoryless first-order Markov model does account well for navigation on the Web. [17] presented the HypTrails method which is used to compare different hypotheses about user navigation in both web [1] and real-life contexts [2].

Folksonomies: The term *folksonomy* was first mentioned by Vander Wal in 2004 on his personal blog.[3] He used this term to describe the underlying structure connecting users, tags, and annotated resources in social tagging systems. While [13] hypothesized that tag distributions follow a power law distribution, thus possibly causing semantic stabilization, [8] showed that after a certain time span, regularities in user activity, tag frequencies and relative frequency proportions could be observed. In turn, this motivated further investigations of tagging systems, especially about the effective extraction of semantically stable content [5] and motivation of tag usage [12]. [9] conducted a user survey on the users' motivation for using social tagging systems, that is, whether users store resources for their own retrieval or social sharing purposes. While there exists a large amount of literature on tagging systems, to the best of our knowledge, there exists only a small amount of work utilizing and analyzing log data from a tagging system. [15] investigated user logs of the social tagging system Dogear, which is internally used at IBM and thus not publicly available, as opposed to BibSonomy. They could find strong evidence for social navigation, that is, users are looking at posts from other people instead of mainly their own. In [7], a thorough study of user behavior in BibSonomy was presented. In the study at hand, we extend these findings, focusing on the actual navigation behavior of users.

3. BACKGROUND

Before we describe and evaluate the hypotheses on navigation in social tagging systems, we recall social tagging systems as well as the formal model of their underlying structure, the so-called folksonomy and briefly introduce our use case BibSonomy. Then, we explain the HypTrails method used for comparing the hypotheses.

3.1 Folksonomies and Tagging

Social tagging systems have established themselves as popular means for organizing and managing digital resources on the Web. The basic idea of a tagging system is that each user can post resources and annotate these resources with freely chosen keywords (tags). By allowing users to assign arbitrary keywords to a resource, they form a powerful alternative to more traditional resource directories or catalogs with fixed taxonomies.

The structure emerging from tagging activities is called a *folksonomy*. [10] models a folksonomy as a quadruple $\mathbb{F} := (U, T, R, Y)$, where U, T, and R are the finite sets of all users, tags, and resources, respectively. The set of *tag assignments* $Y \subseteq U \times T \times R$ is a ternary relation between these sets. Hereby, $(u, t, r) \in Y$ means that user u has annotated resource r with tag t. A *post* from a user u with a posted resource r and the annotated set of tags T_{ur} is defined as a set $P_{u,r} = \{(u, t, r) | t \in T_{ur}\} \subseteq Y$. This also implies that users cannot assign the same tag to a resource twice.

An example for a folksonomy is BibSonomy, a social tagging system for bookmarks and scientific publications (cf. [3]). Note that next to the navigation structure described above, real world implementations often introduce additional navigational features such as showing *related* tags on tag pages or a menu with links to a logged-in user's own pages. Because of this, navigation does not always strictly follow the folksonomy-induced link structure.

3.2 The HypTrails Method

HypTrails [17] is an approach utilizing first-order Markov chain models (MC) and Bayesian inference for expressing and comparing hypotheses about human trails. In our case, the states of the MC are pages in a social tagging system and hypotheses represent ideas about how users navigate these pages. Specifically, hypotheses are formulated as transition probabilities. That is, given a page $s_i \in S$, we define the probability to choose any other page $s_j \in S$: $P(s_j|s_i)$. We introduce several such hypotheses in Section 4. There, we use transition functions $\bar{P}$ to represent hypotheses, which can be converted into the required probability distributions by normalizing the values for each source state s_i.

To obtain insights into the relative plausibility of a set of hypotheses $\mathbf{H} = \{H_1, \ldots, H_n\}$ given data D, HypTrails resorts to the Bayes factor, which compares the marginal likelihoods $P(D|H_i) = \int P(D|\theta)P(\theta|H_i)d\theta$, also called *evidences*, of the different hypotheses H_i. Generally, a hypothesis H_i is more "plausible" than another hypothesis H_j if its marginal likelihood $P(D|H_i)$ is greater.[4] Hypotheses are encoded into the marginal likelihood via the prior $P(\theta|H_i)$ by eliciting the parameters of a Dirichlet distribution from the corresponding transition probabilities of hypothesis H_i and a concentration factor K. The higher we set K, the more we believe in a given hypothesis. This means that the higher K, the stronger we believe in the actual transition probability distribution specified by the hypothesis. With lower values of K, the Dirichlet prior also assigns probability mass to other probability distributions similar to the original one, thus, we give the hypothesis some "tolerance". To understand the difference of our hypotheses in detail, we compare hypotheses based on different values of K. In this work, we express evidences on a log scale. We present corresponding results in Section 6. Note that, while HypTrails gives significant results regarding the ordering of the plausibility of different hypotheses, it is not an absolute measure. We can thus only give a qualified statement regarding the qualitative, not the quantitative difference in performance between the evidence of two hypotheses.

4. HYPOTHESES ON NAVIGATION IN SOCIAL TAGGING SYSTEMS

To be able to compare abstract ideas about how users navigate social tagging systems, we formulate them as transition hypotheses, so we can calculate their evidence with the HypTrails method.

4.1 Basic Hypotheses

First, we formulate basic hypotheses, each representing only a basic aspect of navigation.

Uniform Hypothesis: The uniform hypothesis serves as the baseline for all other hypotheses. It models the assumption that users randomly choose an arbitrary page to visit next, without regard for

[3]http://vanderwal.net/random/category.php?cat=153

[4]According to the reference table in [11], all differences reported in this work are decisive.

the underlying link structure (also called teleportation). Formally, this is expressed as $\bar{P}_{\text{uniform}}(s_j|s_i) = 1$. Since this hypothesis does not require any additional information, it can be considered the least informative one. We expect that any "real" hypothesis capturing a structurally interesting aspect of user behavior, will exhibit a higher evidence than this simple hypothesis.

Page Consistent Hypothesis: Results found by [7] motivate the idea that users often make a transition from a page to itself. This might be accounted for by various reasons, for example to follow pagination, that is, showing the next n elements in a truncated list. This hypothesis is formalized as $\bar{P}_{\text{page}}(s_j|s_i) = \sigma_{id}(s_i, s_j)$.[5]

Category Consistent Hypothesis: [7] found that transitions between two pages often occur between pages of the same category, i.e., after a user has visited a *tag* page, the next page is likely to be a *tag* page again. The same holds for *resource* and *user* pages. The classification of pages into one of these categories is described in Section 5. Thus, the hypothesis states that users stick to the same category. It is defined as $\bar{P}_{\text{cat}}(s_j|s_i) = \sigma_{id}(cat(s_i), cat(s_j))$, where $cat(s_k)$ denotes the category of page s_k.

User Consistent Hypothesis: Similarly to the category consistent hypothesis, this hypothesis assumes that a transition's target and source page belong to the same user. The motivating intuition for this hypothesis is that visitors, who are interested in the work of a specific user, will not only read one, but several of her articles and try to further explore her personomy (i.e., the subset of the folksonomy that only contains the user and her posted resources and tags). It is defined as $\bar{P}_{\text{user}}(s_j|s_i) = \sigma_{id}(user(s_i), user(s_j))$, where $user(s_k)$ denotes the user associated with page s_k.

Folksonomy Consistent Hypothesis: Social tagging systems map links of the underlying folksonomy to actual hyperlinks of the system. For example, the page of a resource contains hyperlinks leading to the page of the resource's owner as well as to the pages of the assigned tags. For that reason, this hypothesis assumes that users navigate only to pages which are reachable in the folksonomy structure and related-tags relations. To calculate reachability, we construct the page graph from the tag-assignments in the folksonomy dataset and (since they are an integral part of the BibSonomy user interface) we add tag-to-tag relations, when tags occur together at the same post. Formally, we define: $\bar{P}_{\text{folk}}(s_j|s_i) = 1$, if s_j is directly reachable in the folksonomy from s_i, 0 otherwise.

Semantic Navigation Hypothesis: Because the folksonomy structure can be used to extract emergent semantics, we aim to investigate the influence of a potential semantic component in navigation behavior. To compute the similarities between two pages, a page is treated as a document. The set of tags which appear on that page with respective frequencies (see Section 5) is treated as the document's "text", represented as the `TF-IDF` vector v_k. The similarity of two pages is then calculated with the cosine measure $cossim(v_i, v_j) = \langle \hat{v}_i, \hat{v}_j \rangle$, where $\hat{v}_k$ denotes the normalized vector of v_k. The hypothesis is defined as $\bar{P}_{\text{tfidf}}(s_j|s_i) = cossim(v_i, v_j)$.

4.2 Combining Hypotheses

In order to investigate possible mutual influences between hypotheses, it is also possible to combine them. In the following, we motivate and describe certain combinations.

Folksonomy Consistent & Semantic Navigation Hypothesis: As described earlier, it is a natural assumption that users utilize the folksonomy structure when navigating a social bookmarking system. If the folksonomy does indeed exhibit notable semantic properties, we should be able to see that adding a semantic component to folksonomic navigation improves the evidence of this hypothesis compared to the bare folksonomy navigation hypothesis. We define the hypothesis as $\bar{P}_{\text{folk-tfidf}}(s_j|s_i) := \bar{P}_{\text{folk}}(s_j|s_i) \cdot \bar{P}_{\text{tfidf}}(s_j|s_i)$.

User Consistent & Semantic Navigation Hypothesis: A similar motivation as with folksonomic and semantic navigation arises when we combine user consistent and semantic navigation. Users are usually thematically restricted in their research interests and can thus serve as a good selector for a limited field of topics. Navigation in the user's personomy is expected to show a strong semantic component. This hypothesis is defined as $\bar{P}_{\text{user-tfidf}}(s_j|s_i) := \bar{P}_{\text{user}}(s_j|s_i) \cdot \bar{P}_{\text{tfidf}}(s_j|s_i)$.

User Consistent & Folksonomy Navigation Hypothesis: The intuition behind combining user consistent and folksonomic navigation is that if navigation is mostly user consistent and partially follows folksonomy-induced links, folksonomic navigation on pages from the same user should yield a good model of navigation. We define this hypothesis as $\bar{P}_{\text{folk-user}}(s_j|s_i) := \bar{P}_{\text{user}}(s_j|s_i) \cdot \bar{P}_{\text{folk}}(s_j|s_i)$.

5. DATASETS

The datasets used in this paper are based on web server logs and database contents of BibSonomy. Since in 2012 the login mechanism was modified, which introduced significant changes to the logging infrastructure, we restrict the datasets to data created between the start of BibSonomy in 2006 and the end of 2011. Anonymized datasets of logs and posts are made available to researchers by the BibSonomy team.[6] Because BibSonomy is a popular target for users who bookmark advertisements, the system uses a learning classifier as well as manual classification by the system's administrators to detect spam. In all experiments, we only use data of users that have been classified as nonspammers.

User and Content Dataset: We use the folksonomy data (all nonspammers with their respective resources and tags) from the BibSonomy database. In the considered time frame, 17,932 users were explicitly classified as nonspammers. They created 456,777 bookmark posts and 2,410,844 publication posts using 204,309 distinct tags. Since we need semantic similarity scores between pages for the semantic navigation hypothesis, we consider all tags which have been used at least twice in order to receive more meaningful results by avoiding typos or rarely used words. With this pruning step, we end up with 65,228 distinct tags.

Request Log Dataset: The BibSonomy log files include all HTTP requests to the system (caching is disabled), including common request attributes like IP address, date and referer, as well as a session identifier and a cookie containing the name of the logged-in user. We only considered direct (i.e., no redirected) valid requests, which have been generated by logged-in nonspammers. Both the referer and the target page of a request must be a retrieval page, that is a page that is used to retrieve information (e.g., a resource or a list of resources; we discuss each considered retrieval page type in the next subsection). For the semantic navigation hypothesis, we had to extract the tag cloud representation of each page. Because a successful request does not imply that the requested page contains any content (e.g., a user tried to filter her collection by a tag that she had not used), we only consider requests that yield a non-empty set of tags using the procedure described in the last subsection. The remaining dataset contains 103,415 distinct visited pages. We recorded 327,060 transitions between these pages. 123,452 transitions were self-transitions (i.e., transitions from a page to itself) and 261,300 were own-transitions (i.e., transitions, where the logged in user owns both the source and the target page). One factor responsible for the large number of self-transitions are pagination effects.

[5] $\sigma_{id}(x, y) = 1$ if x is equal to y, 0 otherwise.

[6] http://www.kde.cs.uni-kassel.de/bibsonomy/dumps/

Page Types and Categories: The pages we consider after filtering the request logs can be assigned to exactly one of six page types. These page types can be grouped into three categories, matching the three folksonomic entity types *tag*, *resource*, and *user*. The six page types (with their corresponding categories) are:[7]

`/user/USER` lists all posts of the requested user `USER` (*user*).

`/user/USER/TAG` shows all posts which were tagged with tag `TAG` by user `USER` (*tag*).

`/tag/TAG` lists all resources with the tag `TAG` (*tag*).

`/url/RESOURCE_IDENTIFIER`[8] describes pages of bookmarked web links to the same web page (*resource*).

`/bibtex/RESOURCE_IDENTIFIER` describes pages that show all publication posts, with the same resource contributed by different users (*resource*). Similar to the previous page type.

`/bibtex/RESOURCE_IDENTIFIER/USER` shows all information that the user `USER` added for a specific publication (*resource*).

A details page for bookmarks is not available, clicking a bookmarked link leads to the bookmarked page (which is outside BibSonomy, thus, these requests are not tracked in the log files).

Tag Clouds as Semantic Page Representations: Since each retrieval page in BibSonomy shows a set of posts, we can define a *tag cloud* for each page. Given a page s_k, its tag cloud is defined as the set of tags with their respective frequencies, which are assigned to the posts of this page. For example, the tag cloud of a page showing two posts, one resource tagged with $social$ and web and another resource tagged with $social$, $concept$ and web, would be $tagcloud(s_k) := \{(social, 2), (concept, 1), (web, 2)\}$. The corresponding document-term vector v_k for the page s_k with the above mentioned tags as features would thus yield $v_k := (2, 1, 2)$, which can used to represent the page s_k.

6. RESULTS

In this section, we evaluate hypotheses about how users navigate on BibSonomy. We first compare our hypotheses (cf. Section 4) on the overall request log dataset. Then we analyze behavioral characteristics of different subsets of the data.

6.1 Overall Request Log Dataset

In this section, we first evaluate the basic hypotheses (cf. Section 4.1) and then the combined hypotheses (cf. Section 4.2) on the overall log dataset (cf. Section 5). Figure 1 shows the results.

Basic Hypotheses: All of the basic hypotheses explain the observed transitions better than the baseline (the uniform hyp.) to varying degrees. This indicates that they all introduce at least some structural properties which help to explain the observed transitions.

Besides this fact, there is a clear order of hypotheses: the user consistent hypothesis works best, the semantic and the folksonomy hypotheses are somewhat similarly plausible, followed by the page consistent and the category consistent hypotheses. Many of the observed effects are explainable by the large number of self-transitions in the dataset caused, for example, by pagination (cf. Section 5): 1. The page consistent hypothesis strongly improves on the uniform hypothesis. 2. The category consistent hypothesis is more plausible than the the uniform hypothesis, even though it directly contradicts

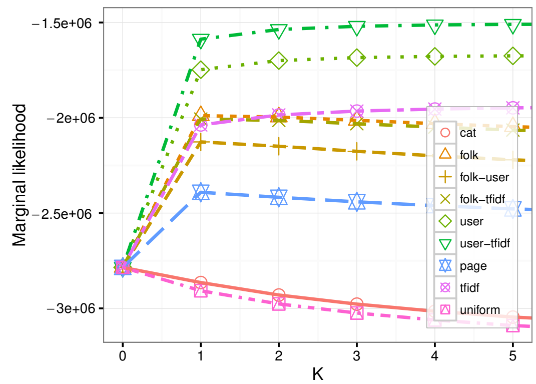

Figure 1: Evidence chart for our hypotheses on the complete request log dataset. Of the basic hypotheses, the user consistent hypothesis explains the data best, followed by the semantic and the folksonomy hypotheses. When combining the user consistent hypothesis with a semantic bias, the evidence improves. This indicates that users are actually semantically biased while navigating through resources. In contrast, combining other hypotheses with the folksonomy hypothesis does not yield better explanations for the observed navigation.

navigation as induced by a folksonomy structure. 3. The user consistent hypothesis as well as semantically induced hypotheses are strongly favored because of self-transitions.

Nevertheless, the user consistent as well as semantically induced hypotheses are also more plausible than the page consistent hypothesis, indicating that their structural properties cover further important factors. That is, the superiority of the user consistent hypothesis indicates that users indeed navigate mostly on their own resources (cf. Section 5). The good performance of the semantic hypotheses indicates that semantic similarity of pages (with regard to tags) is a strong explaining factor for navigation on our dataset.

Finally, we consider the *folksonomy hypothesis* which models the navigation we expect in a folksonomy (see Section 4.1). It performs similarly well as the semantic hypotheses. We observe that the corresponding evidence curve crosses the semantic hypothesis (TF-IDF) for increasing believe factors K. This indicates that the folksonomy hypothesis covers an important factor of the navigation, but fails to model certain transitions, which are covered by the semantic hypothesis. The fact that the folksonomy hypothesis cannot cover certain transitions is due to navigation outside the folksonomy structure as elaborated in Section 5.

Combined Hypotheses: Overall, the combination of the user consistent and the semantic hypotheses performs best, indicating that navigation on BibSonomy can mainly be explained by semantic navigation within the resources of a specific user.

In contrast, combining the folksonomy hypothesis with the semantic hypothesis decreases the observed evidence slightly. Also, combining the folksonomy hypothesis with the user consistent hypothesis decreases the observed evidence dramatically. Both observations indicate that users excessively take advantage of additional navigation features provided by BibSonomy (see Section 3.1) when navigating on their own resources. Interestingly, in Section 6.2 we see that this does not hold when users navigate outside their own scope, that is, on resources exclusively from other users.

6.2 Request Log Subsets

We expect that there are subsets of the data where some hypotheses perform differently than on the overall dataset. Thus, in this section, we investigate different data subsets as listed in Table 1.

[7] Both `/url/RESOURCE_IDENTIFIER` and `/bibtex/RESOURCE_IDENTIFIER` have been restructured and redesigned in mid 2016. They now show combined information about the web page or the publication instead of a list of the same resource.

[8] BibSonomy calculates an identifier for each resource (URL or publication). See http://www.bibsonomy.org/help_en/InterIntraHash for details.

Table 1: Statistics of the request log subsets.

type	source states	links	counts
overall	55,129	149,542	327,060
inside	37,244	105,222	261,300
outside	14,757	28,760	42,193
male	23,090	61,616	130,988
female	5,598	14,413	29,705
neutral	28,726	73,575	161,830
lower_trr	30,368	83,268	176,755
upper_trr	7,084	15,474	32,517
lower_ten	3,459	6,959	15,451
upper_ten	51,542	140,844	307,072
shortterm	10,285	21,912	48,221
longterm	45,535	126,453	274,302

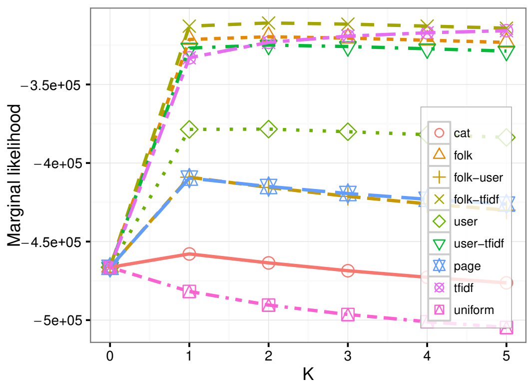

Figure 2: Evidence curves for navigation outside of the user's resources. In contrast to the overall dataset we observe that outside navigation can be explained best by a hypothesis assuming semantic behavior on the folksonomy structure (cf. *folk-tfidf*).

Inside and Outside Navigation: Motivated by the fact that users often navigate on their own pages (cf. Section 5), we investigate whether users behave differently when they are browsing the folksonomy *outside* of their own resources. In particular, we study the transitions where the source as well as the target state do not belong to the browsing user. The results can be seen in Figure 2. Here, the best explanation for the observed navigation is the *folk-tfidf* hypothesis which considers semantic behavior in combination with the structural properties of the folksonomy. This allows us to conclude that while users do not use the folksonomy structure when accessing their own resources (because they most likely explicitly access known publications), they fall back to the folksonomy structure when browsing resources outside of their scope. Furthermore, the evidence for the user consistent hypothesis drops strongly compared to the other hypotheses, because it is restricted to user consistent navigation *outside* of the browsing user's resources. This leads us to believe that browsing outside of the own resources is a process aimed at the discovery of new resources which in turn is not bound to the ownership of resources. Additionally, the fact that the plain user consistent hypothesis performs similarly well as the self-transition hypothesis indicates that the observed user-consistent outside navigation is mostly due to pagination effects.

User Gender: Since gender bias in online systems is an active research area [19], we also investigate the navigation for different genders. In BibSonomy, users can set their gender explicitly. If no gender was set, we assign the label *neutral*, otherwise, we can distinguish between *male* and *female*.

We hardly observe any difference between the genders, thus, to save space, we do not show specific plots. There is only a slight difference when considering the semantic hypotheses compared to the folksonomy hypothesis. It seems that the navigation behavior of male users shows a tendency towards following the folksonomy structure whereas the navigation behavior of female and especially neutral users can be better explained using the semantic hypothesis.

Usage Continuity: Since we expect users to adapt to systems they are using, we investigate if their navigation behavior changes over time. We divide users into *short-term* and *long-term users*, according to the temporal difference of their first and last request. If the difference is less than half a year, we classify a user as a *short-term user* and as a *long-term user* otherwise. In Figure 3, we report the results for short-term users. The results for long-term users are very similar to the results of the overall dataset (cf. Section 4.1).

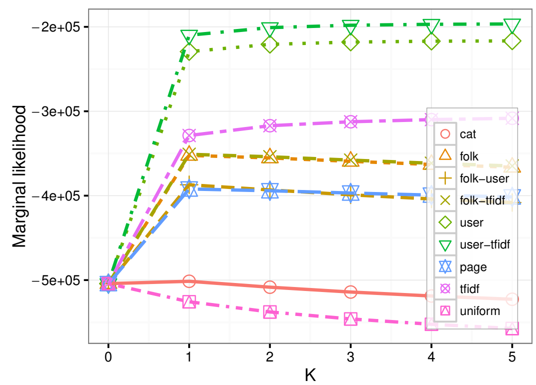

Figure 3: Evidence charts for short-term users ($\leq$ half a year according to their first and last request). We observe a stronger performance of the semantic hypothesis compared to the folksonomy hypothesis and see that the self and folk-user hypotheses perform equally well in explaining navigation. We attribute this to the increased browsing aspect of new users.

When comparing against the overall dataset (or, equivalently, the long-term user group), we observe two conspicuous differences for short-term users. First, the semantic hypothesis performs significantly better when compared to the folksonomy hypothesis. Second, the page-consistent and the folk-user hypotheses are explaining navigation equally well. The former may be explained by the fact that new users are not as tuned to the folksonomy structure as long-term users. Thus, we may actually observe a learning process: the longer users work with the folksonomy, the more they exploit the folksonomy structure in order to navigate their own resources or to discover new ones. The similar evidence curves for the page-consistent and the folk-user hypotheses can be explained by increased pagination effects while exploring the system in combination with the lack of transitions on resources owned by the browsing user. In contrast to outside navigation, here, the lack of transitions on own resources can be explained by the fact that new users have no or a lot less own resources than long-term users.

Tagger Classes: In [12] and [16], different types of folksonomy users were characterized by their tagging behavior. [12] define *categorizers and describers* and [16] identify *generalists and specialists*. Categorizers and describers are classified by their *tag-resource-ratio* (or short *trr*). That is, while categorizers use a small set of different tags for a large number of resources, indicating elaborate category systems, describers use many different tags, indicating a very descriptive approach when tagging. Generalists and specialists are classified using *tag entropy* (or short *ten*), where generalists have a high tag entropy, indicating a wide variety of tagged topics with regard to their resources, while specialists have a low entropy indicating a very specialized set of topics. For both classes we order users according to their *trr* and *ten* values separately and select the upper and lower 30%, respectively. We observe that categorizers and generalists show very similar evidence curves when compared to the overall navigation dataset. For the describers and specialists the curves are very similar to those of short-term users, thus, to save space, we refrain from depicting them.

For both, describers and specialists, we see the same tendency as for short-term users: the semantic hypothesis works better compared to the folk hypothesis and the user-consistent hypothesis has a tendency to perform equally well as the folk-user hypothesis.

The tendency towards semantic navigation over structural navigation on the folksonomy structure can most likely be explained by the nature of the tagging types: Specialists can be considered to be interested in rather few abstract topics, implying a more directed browsing behavior than generalists (whose interests are more varied). Consequently, their navigation is expected to also be more semantically influenced, because of their use of a small, but highly specialized tag subset. As for describers, resources are tagged with more keywords. Thus, for the semantic measure based on TFIDF, calculating the similarity may simply be easier than on the very sparse tagging structures induced by a categorizer's tagging habits.

In general, both types, specialists and describers, can be considered to be of a more explorative nature, as can be seen by the relative performance drop of the folk-user hypothesis and/or the increase of evidence for the self-transition hypothesis.

7. CONCLUSION

Understanding human navigation in web systems is an important step towards improving the design and usability of web pages. In this work, we analyzed navigational behavior of users in a social tagging system. We presented several hypotheses on navigational patterns and evaluated them on a large weblog dataset of the social tagging system BibSonomy.

Beyond confirming the results from [7], that is, that users mainly navigate on their own resources, we were able to show that within these resources, navigation follows a semantic bias (cf. Section 6.1). Also, the semantic hypothesis performs well in general, confirming the semantic component in navigation behavior on BibSonomy.

Furthermore, we studied different navigation subsets and were able to find significant differences in behavior. This includes that even though semantic, user consistent navigation represents a major aspect of the navigational characteristics of BibSonomy, users fall back to the folksonomy structure when browsing outside of their own pages (cf. Section 6.2). Also while different genders did not exhibit interesting behavioral deviations, short-term users, as well as different tagging types, follow certain behavioral patterns matching their individual characteristics. In particular, while it was only hypothesized in prior work [12] that categorizers and describers (as well as generalists and specialists) differ in navigation behavior, we have found specific components of their behavior which differ significantly, thus, indicating that navigation behavior and tagging pragmatics are indeed connected.

Overall, we were able to gain new insights into the underlying processes of navigation in tagging systems, which can be extended and leveraged in the future, for example, by considering new hypotheses, improving navigation experience or extracting semantics.

Acknowledgements: This work was partially funded by the DFG German Science Fund research project "PoSTs II".

8. REFERENCES

[1] Becker, M., Borchert, K., Hirth, M., Mewes, H., Hotho, A., Tran-Gia, P.: MicroTrails: comparing hypotheses about task selection on a crowd sourcing platform. In: Proc. of the 15th iKNOW Conf. (2015)

[2] Becker, M., Singer, P., Lemmerich, F., Hotho, A., Helic, D., Strohmaier, M.: Photowalking the city: comparing hypotheses about urban photo trails on Flickr. In: Proc. of the 7th SocInfo Conf. (2015)

[3] Benz, D., Hotho, A., Jäschke, R., Krause, B., Mitzlaff, F., Schmitz, C., Stumme, G.: The social bookmark and publication management system BibSonomy. VLDB Journal 19 (2010)

[4] Bollen, J., Van de Sompel, H., Hagberg, A., Bettencourt, L., Chute, R., Rodriguez, M.A., Balakireva, L.: Clickstream data yields high-resolution maps of science. PLoS One (2009)

[5] Cattuto, C., Benz, D., Hotho, A., Stumme, G.: Semantic Grounding of Tag Relatedness in Social Bookmarking Systems. In: Proc. of the 7th ISWC (2008)

[6] Chi, E.H., Pirolli, P., Chen, K., Pitkow, J.: Using information scent to model user information needs and actions and the web. In: Proc. of the SIGCHI Conf. on Human Factors in Computing Systems (2001)

[7] Doerfel, S., Zoller, D., Singer, P., Niebler, T., Hotho, A., Strohmaier, M.: What users actually do in a social tagging system: A study of user behavior in BibSonomy. ACM Transactions on the Web 10(2) (2016)

[8] Golder, S.A., Huberman, B.A.: Usage patterns of collaborative tagging systems. JIS 32 (2006)

[9] Heckner, M., Heilemann, M., Wolff, C.: Personal information management vs. resource sharing: Towards a model of information behaviour in social tagging systems. In: Proc. of the 3rd ICWSM (2009)

[10] Hotho, A., Jäschke, R., Schmitz, C., Stumme, G.: Information Retrieval in Folksonomies: Search and Ranking. In: The Semantic Web: Research and Applications (2006)

[11] Kass, R.E., Raftery, A.E.: Bayes factors. Journal of the American Statistical Association (1995)

[12] Körner, C., Benz, D., Strohmaier, M., Hotho, A., Stumme, G.: Stop Thinking, start Tagging - Tag Semantics emerge from Collaborative Verbosity. In: Proc. of the 19th WWW Conf. (2010)

[13] Mathes, A.: Folksonomies - Cooperative Classification and Communication Through Shared Metadata (2004)

[14] Meiss, M., Menczer, F., Fortunato, S., Flammini, A., Vespignani, A.: Ranking web sites with real user traffic. In: Proc of the 1st WSDM (2008)

[15] Millen, D.R., Feinberg, J.: Using social tagging to improve social navigation. In: Workshop on the Social Navigation and Community based Adaptation Technologies (2006)

[16] Niebler, T., Singer, P., Benz, D., Körner, C., Strohmaier, M., Hotho, A.: How tagging pragmatics influence tag sense discovery in social annotation systems. In: Advances in Information Retrieval. Springer Berlin Heidelberg (2013)

[17] Singer, P., Helic, D., Hotho, A., Strohmaier, M.: HypTrails: A bayesian approach for comparing hypotheses about human trails on the Web. In: Proc. of the 24th WWW Conf. (2015)

[18] Singer, P., Helic, D., Taraghi, B., Strohmaier, M.: Detecting Memory and Structure in Human Navigation Patterns Using Markov Chain Models of Varying Order. PLoS ONE 9 (2014)

[19] Wagner, C., Garcia, D., Jadidi, M., Strohmaier, M.: It's a Man's Wikipedia? Assessing Gender Inequality in an Online Encyclopedia (2015)

[20] West, R., Leskovec, J.: Human wayfinding in information networks. In: Proc. of the 21st WWW Conf. (2012)