

The Information Potential and Temporal Elements of Web 2.0 Folksonomy for User Profiling in Personalized Information Retrieval

Wern Han Lim, Saadat M. Alhashmi and Eu-Gen Siew

Monash University, Kuala Lumpur, Malaysia

Abstract

The World Wide Web (WWW) has grown into a rich repository of information. The vast amount of data on the web however would require the users to be skilled in querying generic information retrieval (GIR) systems to meet their information need. This is further complicated by various noises on the web such as spam and advertisement. Personalized information retrieval (PIR) systems have the potential to meet users' information need effectively and efficiently. This paper focuses on the user profiling (UP) aspect of PIR systems as it would directly determine the system's ability to address the users' information need. With the high adoption of web 2.0 systems among the users of the web, web 2.0 systems proved to be an important source of user information for improvement in user profiling. One of such systems is collaborative tagging systems otherwise known as folksonomy. This paper explores the information potential of folksonomy systems in improving user profilers. A case study of the Delicious social bookmarking system was conducted to explore temporal elements of folksonomy which is neglected in previous approaches to improve the performance of user profilers. We conclude that folksonomy systems have the information potential to enhance the performance of user profilers.

Keywords: Information Retrieval, Personalization, User Profiling, Folksonomy, Temporal, Delicious

Introduction

Users constantly find themselves in need for information in today's world; the World Wide Web (WWW) is the biggest repository of information available to the public. The vast amount of information on the WWW however makes the retrieval of particular information in the need of the users difficult. Thus, generic information retrieval (GIR) systems have been developed to assist the users. GIR systems were extremely successful in meeting the users' information need. The ever increasing amount of information coupled with the high amount of noise on the web however would require the users to have IR skills in obtaining the required information such as query formulation as

well as some knowledge in the context of their information need. Thus, the performance of the GIR systems in meeting the users' information need would vary depending on the users and their expertise.

Such circumstances motivate researchers to look for an alternative to GIR systems - personalized information retrieval (PIR) systems. Unlike GIR systems, PIR systems would try and provide the information that the users need without asking for it implicitly. This is done through the PIR system's knowledge about the users as well as the ability to determine the context of the users' information need. Thus, the performance of PIR systems is less dependent on the expertise of the users

while meeting the users' information need effectively and efficiently. Besides that, the IR system must also deal with the potential vocabulary problem such as homonyms and synonyms.

For example, a user with low expertise in IR could produce a query with the keyword 'jaguar' when he/ she as a science student is looking for information about the animal. A GIR system would return results of both the animal as well as the automotive manufacturer. In return, the user would then need to look through the results for the information that he needs. Such situation would not occur within a PIR system. The PIR system would know that the user himself is a science student and the context of the 'jaguar' keyword is in that of the animal and not the automotive manufacturer. Thus, the PIR system would return the suitable results for the users despite the ambiguous query given by the user.

To do so, the PIR system must have knowledge about the users - data about the users collected and inferred into information stored via the PIR system's user profiler. The best source of information about the users would come explicitly from the users themselves but studies found that users are unwilling to provide information about themselves. Thus, implicit collection of user data is the choice of most PIR systems and to do so, the PIR system must have reliable sources of user information. This has been a concern in various researches and with the advent of web 2.0; PIR systems have a credible alternative source of user information. The interaction between users and web 2.0 systems would provide valuable information about the users to be used in the web2PIR's user profiling.

In this paper, we introduce the personalization of IR systems over generic systems to meet the information need of the users in section 2. We then continue on the addition of web 2.0 systems into PIR systems and how web 2.0 could improve user profiling in section 3 and 4. The paper

focuses on the folksonomy component of web 2.0 systems. Our study into this topic is supported by a case study into the Delicious social bookmarking system where we showcase the existence of temporal properties for folksonomy (section 3.3). The temporal aspect of folksonomy is often neglected by past researchers and we look into the information potential from such elements. We summarize our findings and some applications for the improved user profiling in section 5.

Personalized Information Retrieval (PIR)

Most of the information retrieval (IR) systems of the current age such as search engines (Yahoo!, Google etc.) are generic (GIR) - for a given query or request, the same results would be returned regardless of the context or users' interest. An example would be the jaguar analogy given above in the introduction. While the results may change overtime according to the algorithm, the results returned are often generic of various contexts. There is no doubt that GIR systems are successful over the years in meeting the information need of the users. The efficiency of GIR systems is however decreasing with the increasing amount of information and resources on the web. The diversity of such resources together with the associated vocabulary problem in natural language processing (NLP) would require the users to formulate suitable queries to meet their information need. Search query is usually an approximation of the user's information need due to ambiguity (vocabulary problem), mismatches as well as being dependent on the user's vocabulary and knowledge (Wang and Davison, 2008). Thus, a personalized information retrieval (PIR) system is proposed as the solution. While users are accustomed to GIR, it is found that more than 80% of users would prefer to receive such personalized results over generic ones (Lee et al., 2005).

The objective of PIR systems is to provide users with the information that they want or need without asking them explicitly

(Mulvenna et al., 2000). Past researchers have conducted studies into this and found PIR systems to outperform GIR systems various approaches: -

- Re-ranking of results (Daoud et al., 2009, Haveliwala, 2002)
- Recommendation of information (Mobasher et al., 2000, Forsati et al., 2009)
- Query enhancement/ reformulation/ recommendation (Liu et al., 2004, Sieg et al., 2004)

Traditional PIR systems as stated above rely on direct user interaction to learn about the users such as the resources which are viewed, saved, bookmarked and bookmarked by the users especially on the client side. Thus, the PIR system would need to process these resources for to infer the interest of the users. Such resources lack valuable metadata as they are only stated by the authors and thus would require classification of the resources and the content itself to infer the interest of the users for user profiling. Due to the dynamic nature of web resources, classification of such resources is often complex with low precision. Such scenario motivates the research into an alternative source for user information with reliable information and easy processing for user profiling – web 2.0. The improvement in user profiling would directly impact the performance of PIR systems.

The Web 2.0 Approach

The web 2.0 is the next generation of the World Wide Web (WWW). It is an interactive platform for information sharing and collaboration centered on the users known as user generated content and not limited to the authors as the generation before. Web 2.0 platforms include Wikipedia, social tagging (Folksonomy), blogs and social networks. Information on web 2.0 is constantly updated at a high rate with the interaction of users. Researchers in their studies discover and acknowledge web 2.0 platforms as a valuable source of information and their potential to be used within information systems (Marlow et al., 2006) especially in user profiling.

User Profiling with Folksonomy

Collaborative tagging system or otherwise known as a folksonomy system is a system which allows the users and web communities to attach terms, keywords, tags or annotations to shared resources (Golder and Huberman, 2005). Such attachments are short/ brief descriptions which sum up the associated resources (Marlow et al., 2006, Heymann et al., 2008) and help to classify as well as organize resources (Wal, 2005) within their platform. Such addition to resources on the WWW would enhance the metadata of resources which are beneficial to information systems (IS).

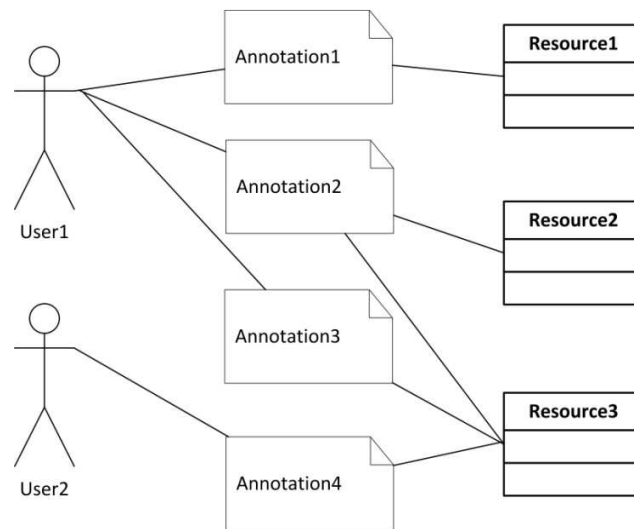


Fig 1. The Folksonomy Model

The basic folksonomy model is built from the 3 main elements: the users of folksonomy systems, the resources on folksonomy systems and the annotations/ tags which are used by the users to describe/ organize folksonomy resources (see figure 1). Thus, the model could be viewed as a tripartite graph with the 3 elements with annotations connecting the users with resources (Mika, 2005, Marlow et al., 2006).

Folksonomy Annotations

In general, studies found that the usage of annotations is varied but highly popular, frequent and stable (Golder and Huberman, 2005). Among the known usage of annotations include: -

- The classification of resource type and is directly used for the personal organization of resources. Classification using annotations were found to successfully replace the use of traditional categorization of classes such as ODP (Xu et al., 2008) for PIR due to their dynamic nature and larger coverage.
- Self noting of the users on their perception or rather a personal metadata about the resources (including quality)

Studies found high overlapping of terms between expert reviews and annotations - 73.01% term overlapping between the content of expert reviews with tags of music resources (Last.FM) (Bischoff et al., 2008). The same observation is made for other resource formats such as the bookmarked webpages of Delicious. Annotations are found to have a high overlapping of terms with both the title and content of resources (Heymann et al., 2008). We could conclude that annotations are highly beneficial in describing the resource while providing additional evidence of resource popularity (Amitay et al., 2009).

Annotations could be used directly as queries due to their high overlapping with queries (Bischoff et al., 2008, Amitay et al., 2009). Beside the semantic meanings which annotations carry, annotations are also found to be a potential measure of the popularity of resources as a form of rating/ voting for tagged resources (Liang et al., 2008). As annotations are assigned by the users, it could reflect the interest of the users more accurately especially in semantics (Han et al., 2010).

Thus, social annotations provide accurate information about the users to be used for

user profiling especially when the annotations are assigned by the users themselves. There is no need for much extra processing of annotations as they are usually short and precise unlike full text methods. Thus, user profiling using annotations are effective and efficient.

Resource Annotations

Traditional PIR systems perform user profiling based on the annotations used by the users directly (Ha et al., 2007, Diederich and Iofciu, 2006) known as 'user annotations' (UA) and not the other annotations which are associated with the related resources, the 'resource annotations' (RA). This is a concept explored by the authors who observed that resources could be used to indicate the interest of the users as users do not only use the annotations which they assigned but also the other annotations of the assigned resources (Cai et al., 2010). Annotation frequency is highly explored in the past to obtain the interest of the users (Noll and Meinel, 2007). The addition of annotations from related resources improves the performance of PIR systems as opposed to non-personalized methods, using only user annotations alone or using only resource annotations alone (Xu et al., 2008). This could be partly explained by the additional information (annotations) which helps the PIR system to establish the context and classification of resources.

The addition of resource annotations allow the PIR systems to profile users who have interacted with those resources and not only users who only contributed (via user annotations) within folksonomy systems. Besides that, these resources could very well appear in other platforms such as social networks which could be interacted by other users and not only via folksonomy systems. This would allow a wider range of users to enjoy the benefits of personalized information retrieval.

Relation between Annotations

A tag cloud could be represented as a vector of annotations weighted according to the frequency of usage by the user (Szomszor et al., 2007, Noll and Meinel, 2007). The information provided by tag clouds (and their direct usage as user profile) are however insufficient due to the following reasons (Michlmayr and Cayzer, 2007): -

- Tag clouds do not maintain the relationship between annotations which hold valuable semantic value and context in the understanding of the users and their interest (Milicevic et al., 2010) especially in dealing with the vocabulary problem especially synonyms and homonyms.
- The popular annotations which are highly weighted within tag clouds are usually very general; thus the specificity of the user interest.
- Does not take into account the temporal element associated with bookmarked resources and annotations.
- Resources might not contain matching annotations with the user profile but are of interest to the user. Thus, tag cloud alone is insufficient as the user profile (but rather need to be used to obtain the interest of the user).

PIR systems based upon folksonomy systems should always maintain the semantic relation between annotations. If two annotations are used in combination with each other (thus co-occur) by a certain user, there is some kind of semantic relationship between them (Michlmayr et al., 2007). The more often annotations are used in association with each other, the stronger the relationship is. Such relationship would make sense to the annotator and thus is not community driven which is usable by PIR systems for user profiling. It is found that 92% of resources within Delicious are annotated with at least 2 annotations (Michlmayr and Cayzer, 2007).

The relationship between the annotations could also be used to obtain the context of the annotations in obtaining the user's interest. Without the relation between annotations, this would not be possible due to the vocabulary problem associated with terms/ keywords such as synonym (group of terms with the same meaning) and homonyms (same spelling of terms with different meaning).

Temporal Element of Annotations

Besides that, annotations do display temporal properties as shown within the adaptive approach such as the Add-A-Tag algorithm (Michlmayr and Cayzer, 2007). Within such approach, the PIR systems take into account the fact that bookmarks (and indirectly annotations) have age and thus signifies the changes in user interest. With this in mind, the user profile should be updated overtime. Under such approach, the users' annotations could be treated as a continuous stream of information and thus sequence could be used to update the user profile. The user profile could be updated in various ways such as via the ant algorithm (an extension on the evaporation technique) where the edges would degrade overtime according to a certain percentage.

The temporal properties of annotations are essential in user profiling to detect the latter interest of the user through newer combination of annotations and bookmarks which are currently more important to the user. In their user study (Michlmayr and Cayzer, 2007) in comparing the co-occurrence approach with the add-a-tag approach, the authors found that users are fond of the performance of both approaches valuing the long term relationship between annotations as well as the ability of the Add-A-Tag algorithm to adapt to recent changes.

User Modeling with Folksonomy

The information potential of annotations could be used for user profiling. The annotations associated with the users are

known as the user's personomy (Hotho et al., 2006). In general, these are the following ways to obtain related annotations of the users according to user action in associating with these resources: -

- Annotations added manually by the users to their profile.
- Annotations used directly by the users.
- Annotations from resources related to the users through users' actions (browsing, viewing, saving, bookmarking, tagging, sharing etc).

Traditionally, the simplest approach is to maintain the annotations in their original form without any processing within the user profile. Such approach, or otherwise known as the naïve approach, models the user profile as a weighted vector of tags/ annotations (Szomszor et al., 2007, Noll and Meinel, 2007) with the weights determined by the frequency of the annotation within the user's personomy. The naïve approach in modeling the user profile is simple, fast and could be used to determine the suitable resources by comparing the cosine similarity of the vector between the user and the resource (Diederich and Iofciu, 2006).

The naïve approach would be suitable for users with only a single interest which is however highly unlikely from the study in the distribution of annotations within the Delicious social bookmarking system (Au et al., 2008). In that study, the authors conclude that users of Delicious have multiple interests and thus proposed to obtain the interest of the users through clustering of resources into multiple annotation vectors and thus obtain the users' interests as topics.

Besides that, the naïve approach does not take into account the relationship between annotations (Milicevic et al., 2010); and thus the context of the annotation which is vital when dealing with the vocabulary problem especially synonyms and homonyms. If two annotations are used in combination with

each other (thus co-occur) by a certain user, there is some kind of semantic relationship between them (Michlmayr et al., 2007). A simple solution to this would be to maintain the user profile as multiple tag clouds (each from the associated resource) (Szomszor et al., 2007) or weighted graph (Michlmayr et al., 2007). The weighted graph approach has the advantage in terms of updating the user profile due to the temporal element which would affect the interest of the users.

Temporal Effect of Folksonomy Resources

Unlike traditional web resources however, resources within web 2.0 systems are enhanced overtime through the users' interaction showcasing the temporal properties of folksonomy systems. This is often overlooked or unexplored by researchers where past researches only take into account the information on folksonomy resources only at the point of interaction. Any information after the point of interaction by the users is not taken into consideration for user profiling. As folksonomy resources exhibit temporal properties which we would prove via our case study, the metadata of the resources increases; which would provide valuable information about the users which interacted with them. Such information include additional annotations (might include temporal based information such as trends etc), new relations between annotations as well as user comments/ notes. This creates an information gap between the user's current point of interaction (information up to this point is being used for user profiling) and subsequent interaction points of other users (new information added by the other users). As such, such information should be used to update the profile of the users as the resources are being updated.

Case Study: Delicious

Delicious is one of the most successful collaborative tagging systems on the WWW for web resources allowing users to bookmark web resources and organize them

through annotations. Thus, we attempt to study the temporal effect of folksonomy systems through a case study of Delicious (see figure 2 to figure 5).

These are the following hypotheses which we would like to assure from our case study: -

- *Hypothesis 1 (H1):* Folksonomy resources do exhibit temporal properties with varying lifespan based upon the interaction of the users
- *Hypothesis 2 (H2):* Folksonomy resources are enhanced overtime through users interaction instead of only on the first day where bookmarks are added and shared on Delicious

We obtained 32401 unique bookmarks from our crawling of Delicious for our case study. From the data extracted, we found that only 4.7% of the Delicious bookmarks have a lifespan of only 1 during. This exhibits the temporal element of folksonomy resources where a big majority of bookmarks on Delicious are still being interacted and updated by users over time. As stated before, past PIR systems would only extract the information at the point of interaction. Thus, users who interacted with the resources on the first day would miss out on the potential valuable information (to establish context and overcome the vocabulary problem) from the resources with lifespan of more than one day. We found that the Delicious bookmarks have a median of 9 days which means that around 50% of the bookmarks have a lifespan of more than 9 unique days. Our findings support the first hypothesis on folksonomy resources having temporal properties.

To support our 2nd hypothesis, we attempt to show that folksonomy resources do receive new unique annotations (we define unique annotations as annotations which terms had not appeared in earlier days for the same bookmark) throughout their lifespan. We observe that the number of unique annotations decreases as the age of the

bookmark increases. An interesting observation would be on the average number of unique annotations (empty ones are not taken into consideration) where the average of unique annotations is pretty consistent between 2 to 3 annotations. This signifies that there are constant new annotations being added by the users to enhance the metadata of the resources overtime. There are bookmarks with relatively high maximum count of unique annotations throughout the days which could contain valuable information.

This would support our hypothesis that the information on folksonomy resources is enhanced overtime and thus should be taken into consideration during user profiling and the update of user profile post user interaction. This is particularly evident from the maximum number of unique annotations for bookmarks in every day as there could

potentially be valuable information which could be used to establish the context of the resources as well as to overcome the associated vocabulary problem.

With both hypotheses supported from our case study, we propose the user profiler of PIR systems to synchronize dynamically with the resource profiler. Thus, new information within the associated resource profiles could be propagated to the user profiles to increase the precision of the user profiler. While past PIR systems have successfully integrated folksonomy into user profiling, none of them synchronize the updates between the user and the resource profile. Thus, valuable information on resources might not be included into the user profile especially for early user interaction. This would be the direction of our next research to determine the actual performance gain of our proposal with the addition of temporal elements.

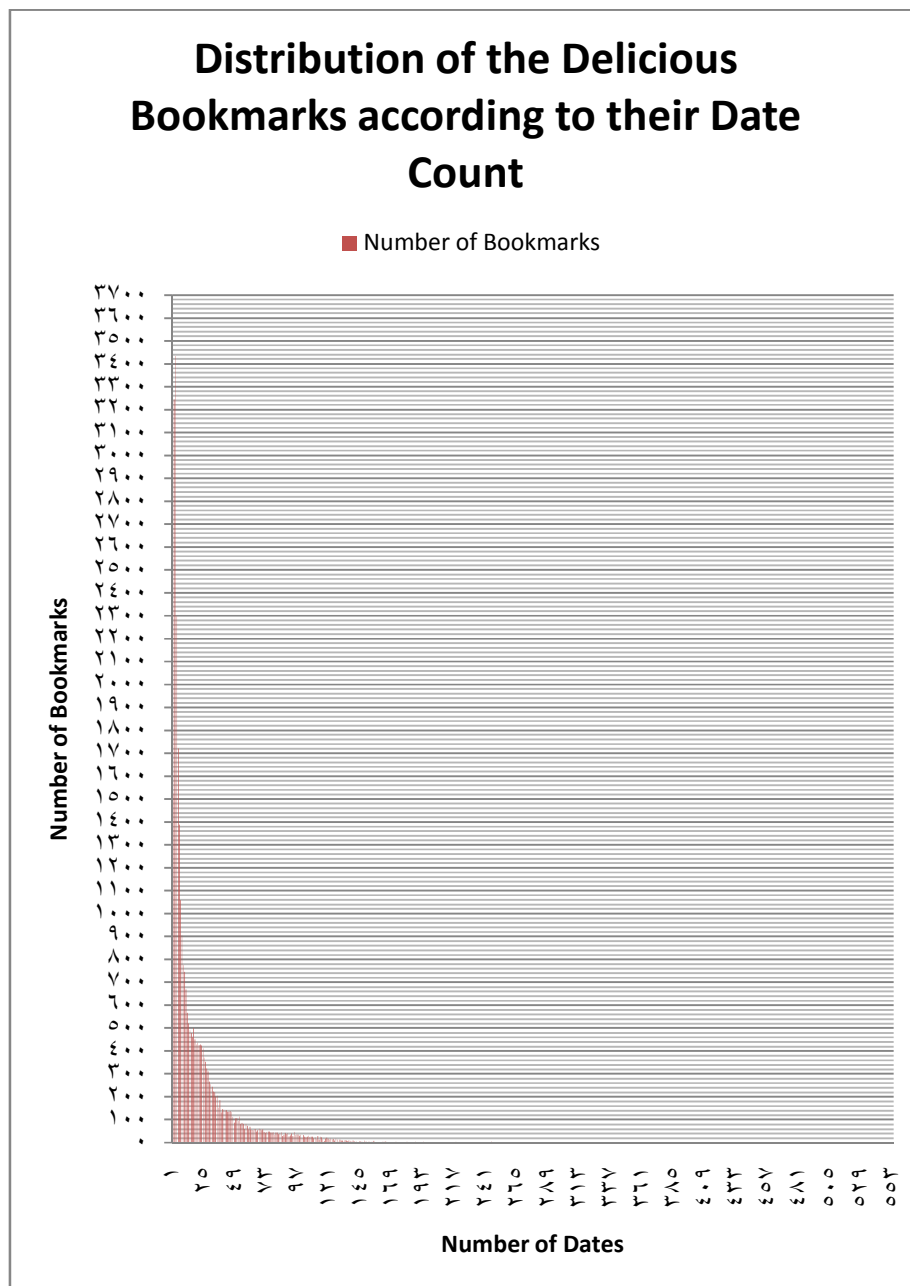


Fig 2. Distribution of the Delicious Bookmarks According to Their Date Count

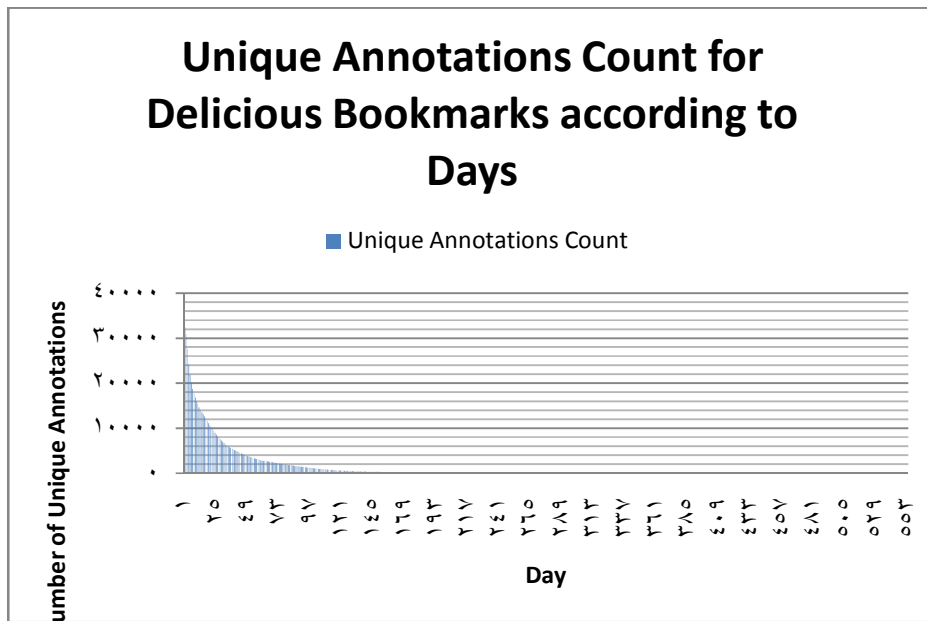


Fig 3. Unique Annotations Count for Delicious Bookmarks According to Days

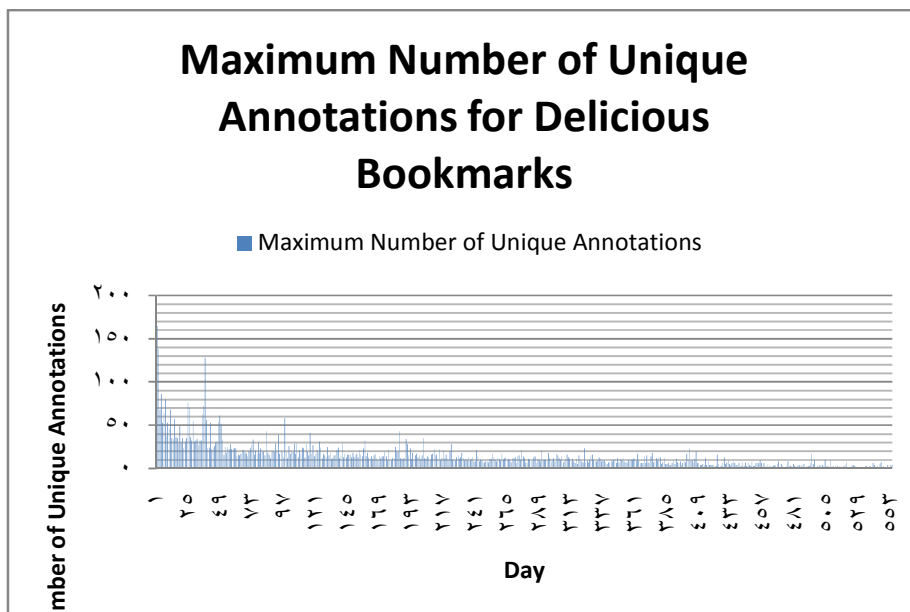


Fig 4. Maximum Number of Unique Annotations for Delicious Bookmarks

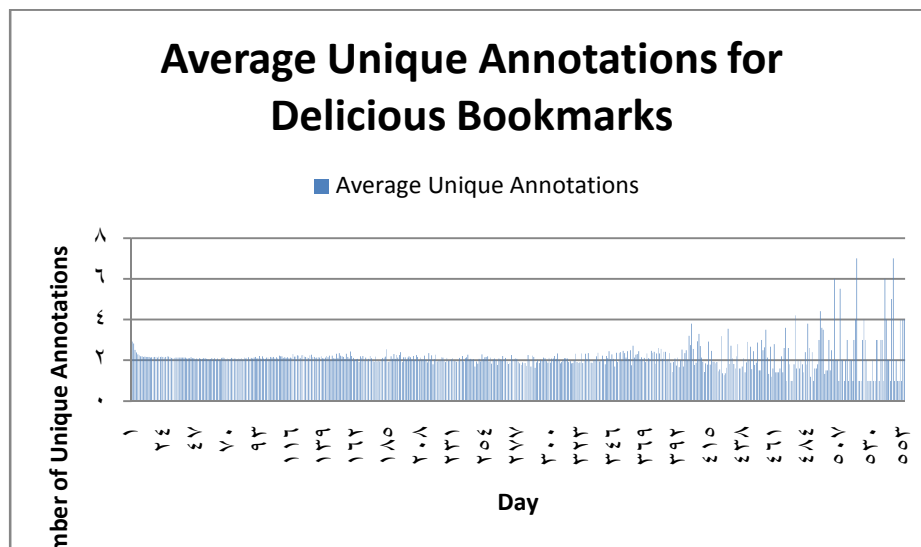


Fig 5. Average Unique Annotations for Delicious Bookmarks

Potential, Impact and Application

Without a doubt that the step into personalized information retrieval (PIR) systems from generic information (GIR) systems is able to meet the information need of the users better. It is seen from our discussion above that most of the users prefer PIR systems over GIR systems and the results from the comparison of various approaches confirm it. The biggest dilemma concerning PIR systems would be on the user profiling – can the PIR system understand the users and their information need correctly? What are the suitable sources of user information to learn about the users? The high adoption rate of web 2.0 systems proved to be a valuable source for user information which would improve the performance of PIR systems in many ways.

Web 2.0 systems do not only act as an accurate source of user information but also allow user profilers to model the multi-interest of users, establishing the interest context (by overcoming the vocabulary problem) and accounting the temporal changes in user interest. Unlike past PIR

systems which would only benefit contributors from the user annotations, the addition of resource annotations would also benefit users which interacted with the resources through various platforms and not only folksonomy systems. The system's confidence and accuracy in the understanding of users would further increase as the information about the resources and their metadata are enhanced overtime by the users of folksonomy systems.

Application of Improved User Profiling

The improved user profiling through web 2.0 systems would benefit information systems (IS) in many ways and not just PIR systems. Within the context of PIR systems, the systems would be able to meet the information need of the users given a user query. From the query of the users, the PIR system would be able to establish the context of the query and overcome the vocabulary problem such as ambiguity by cross-referencing it with information within the user profile. This is held true for any personalized approach such as re-ranking, query enhancements and recommendations.

Within IS, the systems could make use of the user profile in many ways. One of such application is the ability to segment and cluster the users into interest groups using information from the user profile. This would help information systems in identifying suitable user/ interest groups for effective and efficient data mining or information delivery (especially for advertising).

By having the user profile as well, information systems could provide the users with suitable recommendations of information and resources without the need for users to state them explicitly. Coupled with other user-based information such as the location of the user via Global Positioning System (GPS), weather and so forth; information systems could be truly personalized to the users in their daily life.

Conclusion

The addition of web 2.0 systems without a doubt improve the performance of user profilers within personalized information retrieval (PIR) systems which is already outperforming the current generation of generic information retrieval systems (GIR). The main concerns with PIR systems are the system's ability to profile the users which would directly determine the performance of PIR systems. The improvement in user profile would allow PIR systems as well as other information systems (IS) to provide personalized services to a wide range of users and improving the users' satisfaction through such services.

As with our discussion above, web 2.0 systems could very well enhance the user profiler of PIR systems in terms of accuracy, adapting to interest changes as well as modeling the user profile. Past studies have shown the information potential of folksonomy in user profiling. As seen from our case study of Delicious, the temporal elements associated with folksonomy systems create information gaps. This has not been explored in past research within this area. When such gaps are addressed, the

performance of user profilers could be further enhanced especially for users who interacted with the early stages of folksonomy resources before additional information are being added.

The future direction of our research is still focused on the user profiling aspect of PIR systems especially a further look into the information gaps from the temporal elements of web 2.0 systems. We believe that the performance of user profilers could be enhanced by addressing this. We also look at other aspects of web 2.0 systems in improving user profiles. One of such aspects would be the propagation of user information via social networks such as Facebook and Twitter. We would then proceed with our own PIR systems with web 2.0 elements; benchmarking it against the current generation GIR and PIR systems.

References

- Amitay, E., Carmel, D., Har'el, N., Ofek-Koifman, S., Soffer, A., Yogev, S. & Golbandi, N. (2009). "Social Search and Discovery Using a Unified Approach," *Proceedings of the 20th ACM conference on Hypertext and hypermedia. Torino, Italy: ACM.*
- Au, Y. C. M., Gibbins, N. & Shadbolt, N. (2008). "A Study of User Profile Generation from Folksonomie," *Social web and Knowledge Management*
- Bischoff, K., Firan, C. S., Nejdil, W. & Paiu, R. (2008). "Can all Tags be Used for Search?," *Proceeding of the 17th ACM conference on Information and Knowledge Management, Napa Valley, California, USA: ACM.*
- Cai, Y., Li, Q., Xie, H. & Yu, L. (2010). "Personalized Resource Search by Tag-Based User Profile and Resource Profile," In: Chen, L., Triantafillou, P. & Suel, T. (eds.) *Web Information Systems Engineering - WISE 2010. Springer Berlin / Heidelberg.*
- Daoud, M., Boughanem, M. & Tamine-Lechani, L. (2009). "Detecting Session

- Boundaries to Personalize Search Using a Conceptual User Context," *Lecture Notes in Electrical Engineering in Advances in Electrical Engineering and Computational Science*, 39, 471-482.
- Diederich, J. & Iofciu, T. (2006). "Finding Communities of Practice from User Profiles Based on Folksonomies," *Proceedings of the 1st international workshop on building technology enhanced learning solutions for communities of practice (TEL-CoPs'06)*.
- Forsati, R., Meybodi, M. R. & Neiat, A. G. (2009). "Web Page Personalization Based on Weighted Association Rules," *icect in the 2009 International Conference on Electronic Computer Technology*, 130-135.
- Golder, A. & Huberman, A. (2005). 'The Structure of Collaborative Tagging Systems,' *HPL Technical Report*.
- Han, X., Shen, Z., Miao, C. & Luo, X. (2010). "Folksonomy-Based Ontological User Interest Profile Modeling and Its Application in Personalized Search," In: AN, A., LINGRAS, P., PETTY, S. & HUANG, R. (eds.) *Active Media Technology. Springer Berlin / Heidelberg*.
- Haveliwala, T. H. (2002). "Topic-Sensitive PageRank," *International World Wide Web Conference in the Proceedings of the 11th international conference on World Wide Web*, 517-526.
- Heymann, P., Koutrika, G. & Garcia-Molina, H. (2008). "Can Social Bookmarking Improve Web Search?," *Proceedings of the International Conference on Web Search and Web Data Mining, Palo Alto, California, USA: ACM*.
- Hotho, A., Jäschkes, R., Schmitz, C. & Stumme, G. (2006). "BibSonomy: A Social Bookmark and Publication Sharing System," *Proceedings of first conceptual structures tool interoperability workshop*, 87-102.
- Kim, H.-N., Ha, I., Jung, J.-G. & Jo, G.-S. (2007). "User Preference Modeling from Positive Contents for Personalized Recommendation," *LNCS (LNAI)*, 4755, 116-126.
- Lee, U., Liu, Z. & Cho, J. (2005). "Automatic Identification of User Goals in Web Search," *International World Wide Web Conference in the Proceedings of the 14th international conference on World Wide Web*, 391-400.
- Liang, H., Xu, Y., Li, Y. & Nayak, R. (2008). "Collaborative Filtering Recommender Systems Using Tag Information," *IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology*, 59-62.
- Liu, F., Yu, C. & Meng, W. (2004). "Personalized Web Search For Improving Retrieval Effectiveness," *IEEE Transactions on Knowledge and Data Engineering*, 16, 28-40.
- Marlow, C., Naaman, M., Boyd, D. & Davis, M. (2006). "HT06, Tagging Paper, Taxonomy, Flickr, Academic Article, to Read," *Proceedings of the seventeenth conference on Hypertext and hypermedia*. Odense, Denmark: ACM.
- Michlmayr, E. & Cayzer, S. (2007). "Learning User Profiles from Tagging Data and Leveraging them for Personal(ized) Information Access," *Workshop on Tagging and Metadata for Social Information Organization*, 16th Internl. WWW Conf.
- Michlmayr, E., Cayzer, S. & Shabajee, P. (2007). "Add-A-Tag: Learning Adaptive User Profiles from Bookmark Collections," *Proceedings of the 1st international conference on weblogs and social media (ICWSM'06)*.
- Mika, P. (2005). "Ontologies are Us: A Unified Model of Social Networks and Semantics," *Proceedings of the 4th international semantic web conference, ISWC 2005*, 522-536.
- Milicevic, A. K., Nanopoulos, A. & Ivanovic, M. (2010). "Social Tagging in Recommender

Systems: A Survey of the State-of-the-art and Possible Extensions," *Artificial Intelligence Review*, 33, 187-209.

Mobasher, B., Cooley, R. & Srivastava, J. (2000). "Automatic Personalization Based on Web Usage Mining," *Communications of the ACM*, 42, 142-151.

Mulvenna, M. D., Anand, S. S. & Büchner, A. G. (2000). "Personalization on the Net Using Web Mining: Introduction," *Commun. ACM*, 43, 122-125.

Noll, M. G. & Meinel, C. (2007). "Web Search Personalization via Social Bookmarking and Tagging," *Proceedings of 6th international semantic web conference (ISWC) and 2nd Asian semantic web conference (ASWC)*, 4825 of LNCS, 367-380.

Sieg, A., Mobasher, B., Burke, R., Lytinen, S. & Prabu, G. (2004). "Using Concept Hierarchies to Enhance User Queries in Web-based Information Retrieval," *IASTED International*

Conference on Artificial Intelligence and Applications.

Szomszor, M., Cattuto, C., Alani, H., O'Hara, K., Baldassarri, A., Loreto, V. & Servedio, V. (2007). "Folksonomies, the Semantic Web, and Movie Recommendation," *Proceedings of the 4th European semantic web conference, bridging the gap between Semantic web and web 2.0*, 71-84.

Wal, T. V. (2005). 'Folksonomy Definition and Wikipedia,'

Wang, J. & Davison, B. D. (2008). "Explorations in Tag Suggestion and Query Expansion," *Proceeding of the 2008 ACM workshop on Search in social media*. Napa Valley, California, USA: ACM.

Xu, S., Bao, S., Fei, B., Su, Z. & Yu, Y. (2008). "Exploring Folksonomy for Personalized Search," *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*. Singapore, Singapore: ACM.