

# 브라우저 북마크 분류를 키워드로 사용하는 웹페이지 공유를 위한 협동적 URL 태깅 방식

니코 엔카나시온\* · 양현호\*\*

A Collaborative URL Tagging Scheme using Browser Bookmark Categories as Keyword Support for Webpage Sharing

Nico Encarnacion\* · Hyun-Ho Yang\*\*

요약

소셜 태깅 시스템이 직면한 중요한 과제 중 하나는 급격하게 증가하는 태그의 양적 증가와 다양성에 대한 대처방안이다. 구조화된 주석 시스템과는 반대로 태그는 사용자에게 웹 콘텐츠에 주석을 달고 조직화하는 비구조적, 개방적 메커니즘을 제공한다. 본 논문에서는 사용자 정의 태그, URL 키워드, 그리고 분류 폴더 이름을 주요 구성 요소로 하는 폭소노미 기반의 URL 추천 방식을 제안한다. 이 방식은 더욱 개선되어 브라우저의 확장 기능으로 구성될 경우 사용자에게 특정 URL을 분류하는 최상의 방안을 제안할 수 있다.

ABSTRACT

One significant challenge that arises in social tagging systems is the rapid increase in the number and diversity of the tags. As opposed to structured annotation systems, tags provide users an unstructured, open-ended mechanism to annotate and organize web-content. In this paper, we propose a scheme for URL recommendation that is based on a folksonomy which is comprised of user-defined tags, URL-keywords and the category folder name as the major element. This scheme will be further improved and implemented on a browser extension that recommends to users the best way to classify a particular URL.

키워드

Suggestion Algorithm, Bookmark Application, Folksonomy  
제안 알고리즘, 북마크, 응용, 폭소노미

## 1. Introduction

The amount of information that we access in the World Wide Web is fast growing and the problem of organizing all the information that we receive every day becomes more complicated. The info-

mation that we access are normally in the form of webpage and one way to organize these is by using bookmarking sites. In the past decade, bookmarking sites like Pinterest, Reddit, Stumbleupon and Delicious let users save URLs (Uniform Resource Locators) that they find interesting. The

\* 군산대학교 정보통신공학과

\*\* 교신저자(corresponding author) : 군산대학교 정보통신공학과(hhyang@kunsan.ac.kr)

접수일자 : 2013. 10. 30

심사(수정)일자 : 2013. 11. 25

게재확정일자 : 2013. 12. 16

URLs are further organized using keyword tagging.

Even with the advent of bookmarking sites, the simplest form of organizing information (i.e. webpages) is still by using the bookmarking functions offered by the web browsers that we use. Normally, these browsers allow users to organize bookmarks into folders and tag them accordingly. Although some users would comment that the old fashioned browser bookmarks (or “Favorites” as Internet Explorer refers to them) are teetering on the edge of obsolete, it is still perhaps the easiest and most convenient way of saving web pages as soon as landing on one.

Collaborative tagging provides users with navigational cues or “way-finders” for other users to explore information. Social tags are arguably more important in exploratory search and interpretation of information contents by others which provide useful cues for people to discover topics that are relevant. Moreover, the importance of cooperative information searching and result sharing is increasing with the spread of collaborative intellectual endeavors in a network environment. Thus, there is a growing need for effective support of bookmark sharing [1].

The rest of this paper is organized as follows: In Section 2, we discuss some related research projects. Section 3 highlights our motivations for this research work. In Section 4, we state the objectives of the proposed system. We present the system models in Section 5. We present the usage scenario in Section 6. Finally, we conclude the proposal in Section 7.

## II. Related Works

Bookmark sharing can increase the efficiencies of information retrieval and cooperative intelligent activities based on it. Bookmarks have been extensively used in the past and there is abundant

research involving bookmark sharing [1, 4, 5]. A visual bookmark sharing system was proposed in [1] that considers differences in search purposes and user groups. The system extracts useful bookmarks according to the retrieval purpose and shares them within an appropriate user scope. In [4] the researchers discussed the design and implementation of a system that used tagging and shared bookmarks for a social networking application for scientists. Underlying the system is a graph-based data model that links external URLs, system users and descriptive tags. They concluded their paper with a survey of the applicability of clustering and other data mining techniques for the graphs.

Bookmarks can also improve the performance of Web searching. In [3], the researchers proposed combining bookmarks with other features of a search engine to crawl the web for information. The proposed system reflects collaborative evaluation by users by tagging a particular web page with relevant keywords. A user interface was implemented to help store the number of bookmarks as well as the click counts according to the pages saved by multiple users.

## III. Motivation

One significant challenge that arises in social tagging systems is the rapid increase in the number and diversity of the tags. As opposed to structured annotation systems, tags provide users an unstructured, open-ended mechanism to annotate and organize web-content. As users are free to create any tag to describe any resource, it leads to what is referred to as the vocabulary problem [2]. This is why tagging URLs with keywords is not sufficient.

However, when users save bookmarks in their browsers, and organize them using folders, the

names of these folders are more coherent than the tags used. When these folder names are used as additional information to the tags that identify the bookmarks, more meaningful results can be achieved and a more meaningful set of URLs can be suggested to other users.

## VI. Objectives

The objective of this paper is to present a scheme for suggesting bookmarks to users based on the folksonomy of the URLs. The word “folksonomy” is a blend of the words “taxonomy” and “folk”, and stands for conceptual structures created by the people. Related external (i.e. from other users) bookmarks will be recommended to the user depending on the keywords and category used to classify a particular URL. This scheme can be implemented on a browser extension that complements the bookmarking functionality of most popular web browsers (e.g. Google Chrome).

It is important to note that the aim of this scheme is to provide a means for like-minded individuals (e.g. researchers, developers, etc.) to organize information on their browsers while also sharing interesting URLs to their colleagues.

## V. System Model

Social bookmarking and tagging are interesting for Computer Science research because they create usage-driven descriptions of URLs (and potentially any URIs) [4]. In this section, we present how the system data will be structured.

The data set is comprised of the bookmark folder names (which are set as the categories), the user-defined keywords (which are keywords that a user provided as input upon saving a URL), and the URL-keywords which are keywords found on

the URL. This is illustrated in the figure below.

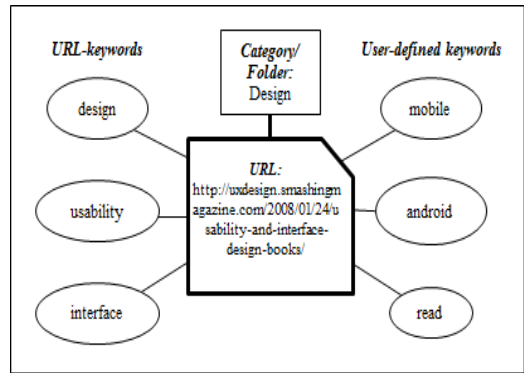


Fig. 1 Data set comprised of the category, user-defined keywords and URL-keywords

In Figure 1 above, the keywords “design”, “usability” and “interface” are retrieved from the URL itself. In addition to these URL-keywords, a user must add more keywords like “mobile”, “android” and “read”, depending on his preference.

### 5.1 Folksonomy

A folksonomy describes the users, resources, and tags, and the user-based assignment of tags to resources [5]. In this research, the URLs are the resources shared among users and are tagged using user-defined keywords and keywords that are found in the URLs themselves.

We present here a formal definition of folksonomies, which is underlying our proposed system.

**Definition 1.** A folksonomy is considered a tuple

$$F := (U, T, R, Y, \alpha)$$

where

- U, T and R are finite sets, whose elements are called users, tags and resources, respectively,
- Y is a ternary relation between them, i. e.,  $Y \subseteq U \times T \times R$ , whose elements are called tag assignments (tas for short), and  $\alpha$  is a user-specific subtag/supertag-relation, i. e.,  $U \times T \times T$ , called is-a relation.

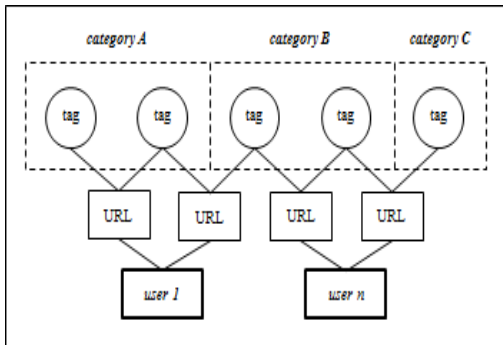


Fig. 2 Data model

In Figure 2, we present the graph-based data model that links the URLs, system users, keywords and categories. A URL will be associated with keywords that all the participants used to tag a particular URL. The keywords are then grouped under 1 or more categories depending on which URL category they were used. It is important to note that each user-URL-keyword-category relationship illustrated above can be represented by the model shown in Figure 1.

The benefit of using these tag set to organize URLs is two-fold. First, the user will be able to efficiently organize his bookmark collection. Second using the tag set, he will get notifications from the system implementing the scheme about other URLs that he might be interested in. The complete description of this recommendation process will be discussed in the next subsection.

## 5.2 Algorithm

Shown in the Figure 3 is the general algorithm for suggesting URLs to a user.

## VI. Usage Scenarios

In this section, we will describe some scenarios where the proposed scheme can be utilized. We assume that the scheme is going to be imple-

mented on a simple browser bookmarking extension or bookmarklet.

```

Step 1. User  $U1$  saves  $URL1$ .
Step 2. Analyse if  $URL1 \in \text{bookmark collection } BC$  .
           if  $URL \in BC$ , inform user.
           else if  $URL \notin BC$ , proceed to Step 3.
Step 3. Retrieve all keywords from  $URL1$ 
Step 4. Prompt  $U1$  to add user-defined keywords
Step 5. Compare  $U1$  keywords set  $k$  to universal set  $K$ 
           (folksonomy)
           where  $k \subseteq K$ 
Step 6. Based on results, suggest similar category
           for  $URL$ 
           if category exists, add  $URL1$  to category
           else create new category  $U1c$ 
Step 7. User chooses  $U1c$ 
Step 8. foreach  $c$  in universal set  $C$ 
           get similarly tagged URLs
Step 9. Suggest top list of URLs to  $U1$  under similar
           categories.
  
```

Fig. 3 URL and category suggestion algorithm

The proposed application will be able to suggest more URLs to a user based on how other users categorized that particular URL. For example, USER 1 bookmarked [www.cnn.com](http://www.cnn.com) (URL 1) and categorized it under “National News”. USER 2, on the other hand, already has another URL, URL 2, bookmarked under the category “National News” and so has USER 2 but under “News”. The application, using the suggestion algorithm, will then recommend to USER 1 other URLs that USER 2 and 3 bookmarked under the “News” and “National News” categories. If more participants were able to bookmark similar URLs under the same categories and tagged with the same keywords, the algorithm will suggest to USER 1 a top list of the those URLs.

There is a need to recognize if a particular URL has already been saved by another user into his browser. This is shown in Step 2 of the URL and

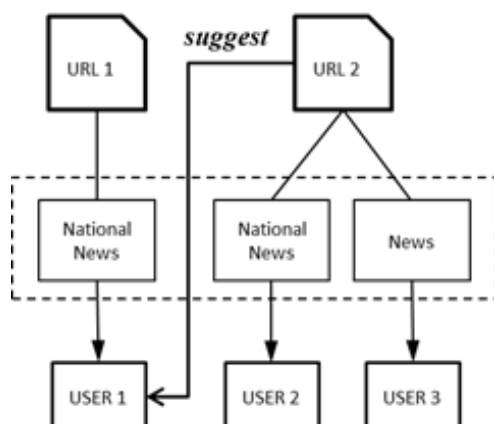


Fig. 4 Suggestion scheme using categories as main elements

category suggestion algorithm presented in Figure 4. A user that is about to bookmark that URL should be able to at least receive recommendations on how to categorize that URL based on what others have used. A user may also be interested in what others saved under a particular folder (category) with the same name as his. This way, information sharing is made more efficient on the browser level even without the user having to access social networking sites (which may be banned in some locations).

## VII. Conclusions

This paper presents a proposed scheme for suggesting bookmarks to users using bookmark categories or folder names as the main element in the folksonomy. This work is motivated by the challenges arising in social tagging systems involving the rapid increase in the number and diversity of the tags. Tags provide users with an unstructured, open-ended mechanism to annotate and organize web content and are sometimes incoherent. We proposed a scheme that attempts to solve this problem by considering bookmark

categories, which are more coherent than user-defined keywords/tags in the knowledge folksonomy. This scheme will be implemented on a simple browser extension that complements the bookmarking functionality of most popular web browsers (e.g. Google Chrome).

## ACKNOWLEDGEMENT

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF-2010-0025460).

## References

- [1] Shoichi Nakamura, Hirokazu Shirai, Hiroaki Kaminaga, Setsuo Yokoyama, Youzou Miyadera, "A Visual Bookmark Sharing System Considering Retrieval Scenes and Its Evaluations", International Conference on Advanced Information Networking and Applications, pp. 730-737, 2007.
- [2] Jin Tan Yang, Wu Lin Lee, Chiu Yen Hwang, "A Web-based Bookmark System with Ontological Approach for Group Content Sharing", IEEE International Conference on Advanced Learning Technologies, pp. 103-105, 2005.
- [3] Kyung Seok Jeong, Hyuk Ro Park, Seok Young Kim, "Improving the Performance of Web Search Using Users' Bookmarks", International Symposium on Information Technology Convergence, pp. 320-324, 2007.
- [4] Marlon E. Pierce, Geoffrey C. Fox, Joshua Rosen, Siddharth Maini, Jong Y. Choi, "Social Networking for Scientists Using Tagging and Shared Bookmarks: a Web 2.0 Application", International Symposium on Collaborative Technologies and Systems, pp. 257-266, 2008.
- [5] Andreas Hotho, Robert Jäschke, Christoph Schmitz, Gerd Stumme, "BibSonomy: A Social Bookmark and Publication Sharing System",

14th International Conference on Conceptual Structures, 2006.

- [6] A. Castagnetti, A. Pegatoquet, C. Belleudy, M. Auguin, "A framework for modeling and simulating energy harvesting WSN nodes with efficient power management policies", EUR-ASIP Journal on Embedded Systems 2012, Oct. 2012.
- [7] A. Swain, R.C. Hansdah, V.K. Chouhan, "An energy aware routing protocol with sleep scheduling for wireless sensor networks", 24th IEEE International Conference on AINA, pp. 933-940, 2010.
- [8] Boon Hee Kim, "Words Recommendation Algorithm for Similarity Connection based on Data Transmutability", The Journal of the Korea Institute of Electronic Communication Sciences, Vol. 8, No. 11, pp. 1719-1724, 2013.
- [9] Boon Hee Kim, "Context Awareness Algorithm for Advanced Usability based on Ubiquitous Sensor Network", The Journal of the Korea Institute of Electronic Communication Sciences, Vol. 8, No. 5, pp. 739-744, 2013.
- [10] Boon Hee Kim, Kim, "Context Awareness Algorithm for Advanced Usability based on Ubiquitous Sensor Network", The Journal of the Korea Institute of Electronic Communication Sciences, Vol. 5, No. 5, pp. 522-527, 2010.



**양현호(Hyun-Ho Yang)**

1986년 광운대학교 전자공학과 졸업 (학사)

1990년 광운대학교 대학원 전자공학과 졸업(석사)

2003년 광주과학기술원 정보통신공학과 졸업(박사)

1989년~1990년 삼성SDS 근무

2005년~현재 군산대학교 정보통신공학과 교수

관심분야 : 무선데이터통신, RFID/USN etc.

저자 소개



**니코 엔카나시온  
(Nico N. Encarnacion)**

2011년 West Visayas State University, Philippines BS in Information Technology

2011년~현재 Department of Information and Telecommunication, Kunsan National University, South Korea, (Master Student)

※ 관심분야 : Wireless Data Communication, Wireless Sensor Network.