

# Online WordNet Based Tagging System for Social Sharing and Retrieval of Images on Visited Pages

Tansel Özyer

Department of Computer Engineering, TOBB Economics and Technology University, Ankara, Turkey

E-mail: ozyer@etu.edu.tr

**Keywords:** social networks, collaborative filtering, web information retrieval, XML

**Received:** October 12, 2008

*With the advent of internet technologies, users get involved in collective organizations that feed from collective intelligence mechanism to make further recommendations, inferences on unseen examples. In this study, as users surf on the internet, they are capable of picking pictures liked from web pages to their favorite list with definitions specific to our system. After images are tagged with definitions, they are filtered with respect to commonality measure to deduce significant definitions for storage. These tags help users to make inquiries and broaden up their search on the tags they are interested. While doing tagging process, all images are shared collaboratively and that is bolstered with Word net.*

*Povzetek: Opisano je sprotno označevanje slik s pomočjo WordNeta v socialnih omrežjih.*

## 1 Introduction

### 1.1 Problem Definition

Researches point out that nowadays news, videos and pictures of the 7% of the sites visited is being tagged [1]. Also, 28% of people benefit from tagging.

While surfing on the internet, users are capable of bookmarking resources such as images on the internet. In a social environment people share information and exchange. Instant exchange of ideas and finding the image sought can occur by clicking on an image found at a document and by asking "What are the images similar to that image?" We are looking forward to finding similar images to that in social behavior according to the descriptions made. Otherwise, traditional image retrieval techniques will remain insufficient according to real time requirements at first.

Sites such as deli.ci.ous.com, <http://www.flickr.com> (picture sharing) and <http://www.youtube.com> (video sharing), digg, fark, StumbleUpon, Reddit, Slashdot are some of favourite ones as a social bookmarking system [16]. It has become widespread because it is very simple and practical process. However, there are also some problems with tagging in collective manner such as using synonym, polysemy words and etc. Tagging a picture depends on its interpretation by the owner. Keywords entered for the picture give the bookmarking info. User, himself can organize his own data besides, he can share his information with other users in the community.

For example, frequently used information bookmarked is stored at <http://deli.ci.ous.com> keeps track of popular tag info. It is categorized according to the group name which is the tag value. When someone indicates that a web page is about a topic then the link is being put together with the other links having the same topic specification.

Whenever the user wants to reach the links of a specific topic then, all links that are related are brought back. Throughout the development of web site, some statistical methods turn the web site into more efficient situation. There are some basic rules about formatting and describing tags [2]. When users obey, it becomes easier to access and manipulate [3].

### 1.2 Motivation

Bookmarking can be supported by tagging utilities. Tagging is beneficial for searching bookmarked items. This means, pictures are supported by words/terms for description. Basically, background, action, objects of an image can take part. Stand-alone tagging can be used to some point but support in a collaborative environment for sharing can be useful for finding similar images to an image. One can find or suggest similar alternative images to that image. All tagged information may help users share images between each other. Tags may need to be categorized and tag entries should be unified in a sense that they mean the same meaning to each other while doing similarity check.

### 1.3 Contribution

This study proposes a collaborative infrastructure that helps users store pictures that they have seen on web pages during surfing. Basically, system consists of two modules: One module is an add-on application that will work on Firefox Mozilla browser and the other one is the web page itself. The add-on application helps user tag and store pictures that will remind of something meaningful to him. Tagging is being used very often in variety of internet applications in order to expedite

access to the resources on the web. Another aim of tagging is using internet as a private agenda. Tagging merits users organizing data with respect to their special needs.

Web 2.0 is an emerging platform that helps fast development of interactivity and interconnections of web data. Our study benefits from web 2.0 and social bookmarking utilities.

In this study, we propose a novel study that employs the use of Wordnet in order to unify the tags described by the users and image similarity is measured by using Jaccard coefficient. Tools and technologies that have been used for our system are described at the next section.

## 1.4 Technologies and Tools

C# and ASP.Net have been used to build the web site and database has been used at the server side. Firefox has been used as the web browser tool and communication between web services and Firefox has been provided with AJAX (Asynchronous JavaScript and XML) technology. We have created an interface for add-on application to be used via XUL and JavaScript for functionality. We have used VS Studio and MS SQL Server Express, Firefox and Firebug Extension Developer plug-in for Firefox.

## 2 System overview

As is mentioned, system is formed of two parts. These are Firefox add-on application that works on browser in order to use on the client for tagging and the repository at the server.

### 2.1 Firefox add-on application

Mozilla Firefox is a web browser that is licensed under MPL/GPL/LGPL. As of today, Firefox is being used around 45% (was %25 in 2006) of users and it becomes one of the most common web browser all over the world. One of the main reasons for the gradual use of the web browser is that it's an open source application and third party developers can easily implement plug-ins onto the browser as an extension.

Firefox extensions basically built over two structures. One of them is XUL(XML User-Interface Language) for interface definition; and the other one is JavaScript that has been used for the client side functionality requirements.

XUL is an XML language that has been designed to define new user interfaces or modify existing ones. It is capable of using XSLT, XPath, and DOM standards. Entire user interface of Firefox has also been defined with the same language that makes it convincing about how powerful language it is [4].

Add-on tool listens to the click events on a web page that is currently open. If a click occurs on a page, it detects whether the click is on an image or not. When a user clicks on a picture, the application will display a pop up window accordingly. Next, the user can check his user information and get the potential tagging recommendations in cooperation with the server and the

user can arrange the tags for the picture by also taking to the given recommendations into account. After arranging the tags, tag information about the image will be sent to the server. Later on, tag information that has been processed can be retrieved back from the server again for confirmation to the user.

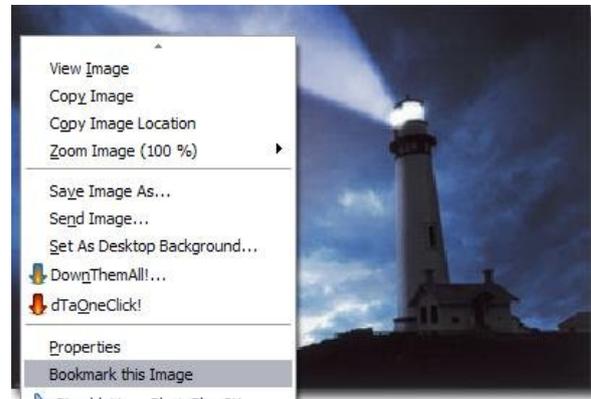


Figure 1: Right Mouse Cursor Click.

During the process, one of the most important points is the arrangement of the obtained tag information. Information can be processed at both client browser and server. However, it will be better to pre-processed at client side by using JavaScript in order to reduce the workload of server. Arrangement of tagging information has two meanings: First one is stripping out the words from whitespace and non-alphanumeric characters; the second is stemming words in order to reduce sparsity and group words under specific categories disregarding the suffixes (e.g. performing, performed, performance and etc.). Popular stemming algorithm, Porter's stemming algorithm has been used [5]. All terms are pre-processed at client browser side and sent to the server.

After cleaning up the words, they are separated with a delimiter. The reason for using delimiter is that it is impossible to send a JavaScript object array directly to a web service. Instead, a string will be sent as a parameter. Another point is to reduce the communication cost as much as possible. In order to accomplish this, recent technology, Ajax has been used to prevent refreshment of entire page at each time a change occurs. This is done by using the basic Ajax object XMLHttpRequest. XMLHttpRequest object provides an asynchronous connection between two parties that is, user is not supposed to wait for the reply of the request sent back and only related portions of the web page is refreshed instead of entire page[8].

An image can be tagged to describe the content of the image for describing the content as well as specific notes taken onto it. Description part of the image content is composed of different parts as <objects>, <actions>, and <background> information fields (Figure 2). Here, a user is free of entering all the information about the picture for tagging at the browser instantly and keep visiting pages during the day.

## 2.2 Server Application

There are two main tasks at the server side. These are processing and storing tag info and forming the infrastructure for sharing environment.



Figure 2: Tagging an Image.

### 2.2.1 Communication with browser add-on

The most prominent job of the server related to the add-on tool is the language processing of tag information. In other words, information is checked. If the entries are adjective, verb, adverb and etc. WordNet database has been put in use to accomplish this at the server side.

WordNet was created and has been maintained by the cognitive Science Lab of the Princeton University. It is a semantic lexicon for English. It is used to group words into sets such as synsets. By employing wordNet, it is possible to access the semantic relations between words and their definitions and roots. [6].

After pre-processing, tagged information is sent to the server and server revises them and sends back to the client browser independent from the user for confirmation. User is free of doing modifications on the tags and after confirmation, server stores user, image link, and tag information at the database.

Another aspect of the communication is the user control. This is performed in the following way:

Firefox add-on application at the client first checks if the user whether signed on or not by holding a variable kept in his DOM-like structure.

If signed on then the user is authorized. Otherwise,

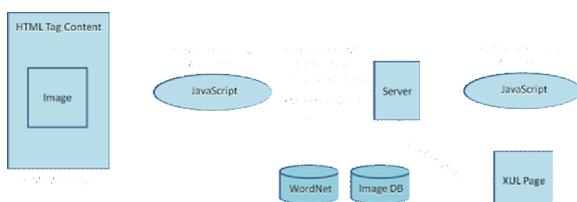


Figure 3: The Flow Diagram of Tagging Process.

user name and password are requested. Ajax based requested user name and password are obtained and sent to a web service with GET method. This service basically fulfils the authentication of the system. Server controls the access with one time login.

The web services are used to support different applications to interwork via machines and exchange messages in XML format [7].

Data flow can be summarized as the following: An image can be accessed via DOM structure. Tag element content surrounding the image. In this study, we extracted the alt and innermost tag element content as the image content information (Figure 4). Non-alphanumeric characters are removed and content is spit into words. These later are put into a string to be able to send to a web server as a parameter. This information is sent to the web service to the server for feedback to the end user. Details of the feedback mechanism will be discussed later but shortly, it includes the interaction between the image db and wordNet in Figure 3.

```

html>
<head>
<title>... </title>
<body>
...

...
</body>
</html>
    
```

Figure 4: HTML Page.

After the tag information is arranged, the feedback is sent to Firefox browser add-on in XML file. Firefox presents it to the end user with an interface that is constructed with XUL.

User is free of accepting all tags specified or they can be changed. After confirmation, final tag information is sent to another web service at the server side. All parameters are resolved and they are used for updating database tables (user id, link address, word/word type).

### 2.2.2 The Database Model

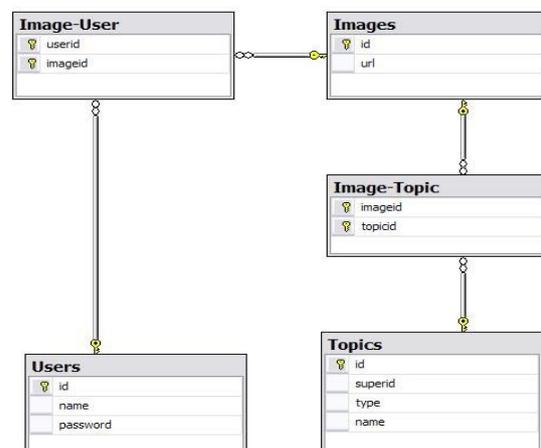


Figure 5: Database Relationship Model.

Database relationship model is given in figure 5. There are five tables they are namely, Users, Image-User, Images, Topics, and Image-Topic tables.

Users table contains the user specific information; Images table contains image specific link information. The most important table is the Topics table. It keeps the tag information: Words, their types (adjective, noun ...) and their hypernym. Topics table is linked to WordNet database.

### 2.2.3 Wordnet

WordNet has been used for tagging words. While tagging, in order to give idea to the end-user and grouping tag content is provided. It contains approximately 147000 words. WordNet has been defined in a relational database and this formation has been used in our system[10]. Synonym set of a word can be used by doing inquiries on the database. The SQL query that has been used for this purpose has been given figure 6.

```
select lemma from word where wordid in
(select distinct wordid from sense where
synsetid in
(select synsetid from sense where wordid in
(select wordid from word where lemma=@word)))
```

Figure 6 – SQL Query for Finding Synonym Set(Synset)

We have used stored procedures. By using the SQL Query in figure 6, synonyms of a word can be retrieved from the database. While tagging words, synonym words are found. Another function of WordNet is grouping the words in terms of their categories. *categorydef* in wordNet[10] can be used for this purpose. Categories are noun, adjective, adverb and verb. These categories are also subgrouped. For example, nouns can be event, location, and animal, human-being nouns. Each word has a *category\_id* and it is related with *synset* table. By using these properties, category of the word can be found(Figure 7).

```
select [name] from categorydef where categoryid in
(select categoryid from synset where synsetid
in
(select synsetid from sense where wordid in
(select wordid from word where
lemma=@word)))
```

Figure 7- SQL Query for Finding Category of Word

Tag information of the image (*alt* property and surrounding innermost element content) are used identify the content of the image.

Whenever potential tag information has been received, each word that has been obtained will be matched with similar words existing at wordNet database and potential tagging content is intended to be increased. Here, type info keeps the type of the word(noun, verb, adverb and etc.) and superid field keeps the broader meaning of word(apple is a fruit).

### 2.2.4 Web Services

One of the major parts of our system is the web services. There are three web services that was used. These are for:

- User authentication
- Feedback mechanism
- Update of database server with confirmed data.

The most important web service is the feedback mechanism. Our aim is to categorize them according to their types. Object (noun), action (verb) and background. After composing a string message having delimiters in between words, for each word, first we check whether the specified word exists in database or not. If exists, type info is obtained from our database content otherwise, for all similar words and their corresponding types are sought. If found then type of words are the same as their similar words' type. Otherwise, type of that word is found directly from wordNet. At this stage, background, action, noun situation are checked. For example, for the word play there are different type info to be obtained. It has the noun type (as game) or action (as to play). In this situation, it ends up with the type of majority (e.g. Five of them are noun and two of them are verbs then type is accepted as noun).

After all, strings acquired by the service for feedback are sent back to browser XUL window with Ajax for confirmation.

### 2.2.5 Website

According to the procedure described above, tag information from the end user is obtained. After this, one further is taken a web site as a sharing medium has been constructed. Web site has been implemented by C# and ASP.Net. Users sign in with their user name and password. After logging on, they can view what they tagged for what images at the web site. There is a search interface for their search at their bookmark page. It provides a search environment according to the tag info described. Search can be done either on user's tagged image set or entire system's bookmarked image set.

Search function can be done by entering words for selected fields (object, action, background). Search results do not only depend on the exact words as search criteria, they depend on the words similar to search



Figure 8: Web Site with Bookmarked Images and Tags.

criteria. Search operation depends on the similarity metric that will be discussed at the next section. Web site has been implemented by using Web user controls [11].

Web site contains the bookmarked pictures and their tag information below. Web site has been supported by the CSS (Cascading Style Sheets) for visual display. Tag information is composed of object (noun), action (verb) and background information of the image.

### 2.2.6 Image Similarity Search

At the web site user is capable of doing search in the local bookmark list as well as the entire site (across users). Image search aims at finding similar images according to the tag criteria specified at the search entry. In the search process, Jaccard similarity is applied. Given a set of words as tag information, we find the jaccard similarity between two images in terms of set of words as,

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \tag{1}$$

Where  $A = \{w_{A1}, w_{A2}, w_{A3}, \dots, w_{An}\}$  and  $B = \{w_{B1}, w_{B2}, w_{B3}, \dots, w_{Bm}\}$

Jaccard similarity will give back a similarity score between 0 and 1. Our Jaccard similarity is modified. It uses WordNet to find the number of synonym words. As the number of synonym words increase, the contribution of that word to similarity result decreases.

This method is the combination of WordNet and Jaccard similarity. For example given tags of image A and image B,

A = {baby, laughing}

B = {baby, sleeping, bed}

Table 1: Words and their # of synonym word.

Words	# of Synonym Words
baby	13
laughing	8
sleeping	11
bed	36

$A \cap B = \{baby\}$  and  
 $A \cup B = \{baby, laughing, sleeping, bed\}$

$$J = \frac{baby}{baby + laughing + sleeping + bed}$$

Baby = 1/13, Laughing = 1/8, Sleeping = 1/11, Bed = 1/36

$$J = \frac{1}{\frac{1}{13} + \frac{1}{8} + \frac{1}{11} + \frac{1}{36}} = 0.56$$

Similarity score between images A and B is 0.56. In addition to image similarity, only words entered as search criteria can be used. For example, instead of all tags, baby or laughing could be entered. The picture below gives how the search criteria is entered.

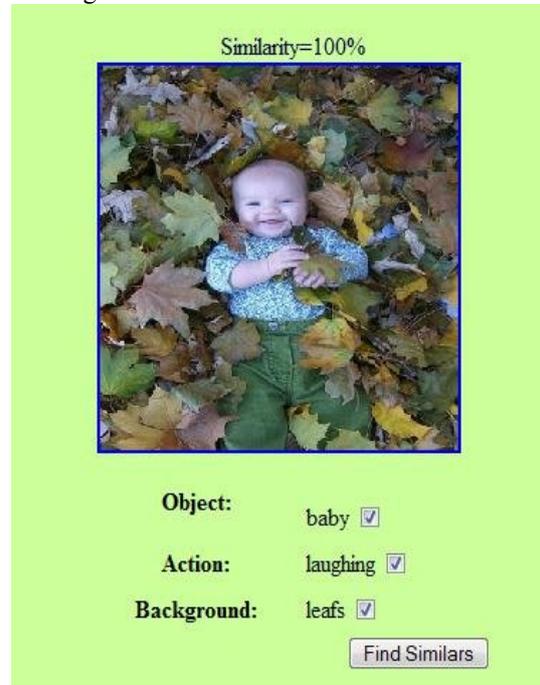


Figure 9: Image Search.

During the search, basically there is an extreme case such as finding the intersection of two sets as 0. For example rose and daisy can be grouped under flowers and hypernym of these words are used instead by taking the  $\beta$ (it is specified as 0.5) to the power of minimum distance they meet up to the tree.

Figure 10 gives the search results for laughing baby criteria entered. It is also possible to search from one of the results. Search results assist user and he can find more results by using one of the images as the query criteria.

### 3 Experiments

We performed experiments over 20 students and they collected images on specific topics given. Topics are



Figure 10: Search Results.

human beings, vehicles, sports, nature, smart phones, computers, house appliances, clothes, fashion. First, we allowed users find their 50 images per user for each category. They tagged their images and before doing our similarity search, we performed association rule mining [15] algorithm to find the patterns frequently approached. We kept the support value as 0.05, 0.01 and 0.15 and we considered frequent patterns over that value for images that are the same (after link information and fast detection of images of possible likeness). We have considered the union of patterns that are subsets of frequent patterns. Besides, images are checked if they are relevant to categories.

Normally in our system, similarity results less than a user specified threshold are ignored. (Default value is 0.1) but in experiments we adopted to idea of taking top  $n$  images from the search result set. We asked users to do five queries for each category and we checked the results if they are related or not.

While doing comparisons, we used f-measure that is dependent on precision and recall values [14].

$$\text{precision} = \frac{\text{no of correctly recommended images}}{\text{no of recommended images}} \quad (2)$$

$$\text{recall} = \frac{\text{no of correctly recommended images}}{\text{no of relevant images}} \quad (3)$$

$$f\text{-measure} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (4)$$

We asked users to create search corpus as 50 words as search criteria from each user (30 objects, 20 verbs and 10 background). We constructed random search criteria by using our corpus for each category (15 queries for each) and we obtained the f-measure results below.

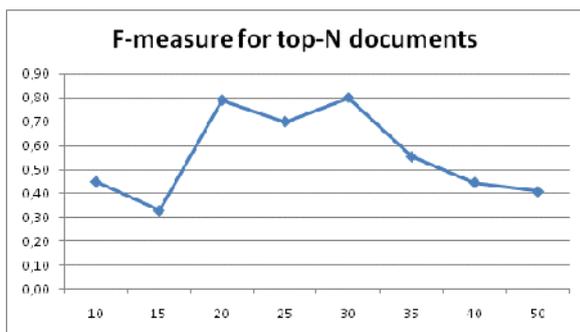


Figure 11: F-measure Value Results for Top-N Documents.

We restricted our queries to top-k queries in order to find the f-measure value. top 20 and top 30 query results in scoring return the highest scores around 80%. We got the lowest in top 15. As the top scoring result increases, f-measure starts to decrease. Our results are promising in terms of top 20 and 30 queries. Here, we also intend to limit the search records for time limitation.

## 4 Conclusion

In this study, we implemented a system that lets users to bookmark pictures they want to keep with their

tag info according to image content. Later on they may reorganize the tags and they may also do search by using tag info. Rather than image content, we tried to keep the tag information in order to do search and find similar images based on the tag field values entered.

In the future, we will also incorporate different utilities of Wordnet such as antonyms, and etc. Also, usage and rating of images will be engaged with heterogeneous scoring mechanism.

## Acknowledgement

I would like to extend my special thanks to anonymous reviewers for their valuable comments and contributions.

## References

- [1] John C. Paolillo and Shashikant Penumarthy(2007), The Social Structure of Tagging Internet Video on del.icio.us Proceedings of the 40th Hawaii International Conference on System Sciences.
- [2] Lee Rainie(2007), Interview: Author David Weinberger Describes How Tagging Changes People's Relationship to Information and Each Other Pew Internet and American Life Project January 31.
- [3] <http://del.icio.us/help/tags>
- [4] <http://www.xulplanet.com/tutorials/whyxul.html>
- [5] M. F. Porter(1980), An algorithm for suffix stripping no. 3, pp 130-137, July.
- [6] <http://wordnet.princeton.edu/>
- [7] <http://www.w3.org/TR/ws-arch/#whatis>
- [8] Nilayan Sharma(2005), Snappier Web Applications with XMLHttpRequest, April.
- [9] Lihui Lei; Zhenhua Duan(2007), "Integrating AJAX and Web Services for Cooperative Image Editing," IT Professional , vol.9, no.3, pp.25-29, May-June.
- [10] Che-Yu Yang; Hung, J.C.(2006); Word Sense Determination using WordNet and Sense Co-occurrence Advanced Information Networking and Applications.
- [11] Microsoft Developer Network, [http://msdn2.microsoft.com/enus/library/3457w616\(VS.71\).aspx](http://msdn2.microsoft.com/enus/library/3457w616(VS.71).aspx)
- [12] <http://www.w3.org/Style/CSS/>
- [13] Sudkamp, T.; Cross, V.(1996), Toward fuzzy-valued matching and similarity assessment, Fuzzy Information Processing Society, NAFIPS. Biennial Conference of the North American
- [14] Gerard Salton(1983), McGill, M., 1983. Introduction to Modern Information Retrieval. McGraw-Hill, New York.
- [15] R. Aggrawal, T. Imielinski, and A. Swami(1993). Mining association rules between sets in large databases. In Proceedings of the ACM-SIGMOD International Conference on Management of Data. Washington, DC.
- [16] List of social bookmarking applications downloaded from: [http://en.wikipedia.org/wiki/List\\_of\\_social\\_software#Social\\_bookmarking](http://en.wikipedia.org/wiki/List_of_social_software#Social_bookmarking) on Feb 20, 2009.