Pertanika J. Sci. & Technol. 24 (2): 493 - 506 (2016)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Automatic Tags Generation in Folksonomy for Learning Resources Reuse and Sharing

Ching Chieh Kiu

School of Computing and IT, Taylor's University, 47500 Subang Jaya, Selangor, Malaysia

ABSTRACT

With the proliferation of Web 2.0 technologies, folksonomy which is also known as social tagging or collaborative tagging is widely used by learners to annotate and categorize their learning resources. In a folksonomy system, the tags are added by learners to the learning resources, hence the tags are often ambiguous, overly personalised and imprecise. In addition, conjugated words, compound words and nonsense words may be used in tagging and shared among a group of learners. This has resulted in an uncontrolled and chaotic set of tagging terms that cause learning resources searching, reuse and sharing to become ineffective. In this paper, we present a content-based approach which automatically generates tags from a learning resource using Part-Of-Speech Tagging and K-Means Clustering techniques. The generated tags are more precise and unambiguous which can improve learning resources searching, reuse and sharing among learners.

Keywords: Folksonomy, clustering, K-Means, Part-of-Speech, collaborative tagging, learning content reuse, learning content Sharing, learning process, automatic tagging

INTRODUCTION

With the increasing popularity of Web 2.0 technologies, folksonomies (a method for collecting, organizing and creating tags) are emerging to allow learners to categorise and annotate learning content through collaborative tagging or social tagging systems. In social tagging systems, learners freely choose their own vocabulary, so-called tags, by adding explicit meaning which may come from inferred understanding of the learning contents. In this way folksonomies provide a new mechanism for learners to retrieve their learning resources via tag clouds (unstructured navigations). In addition, the mechanism also enables the learning resources to be shared and reused with their peers in an e-learning environment. The example of learning resources uploaded and tagged by learners with freely chosen tags

Article history:
Received: 8 December 2015
Accepted: 22 April 2016

E-mail addresses:
kiuchingchieh@gmail.com, chingchieh.kiu@taylors.edu.my
(Ching Chieh Kiu)

ISSN: 0128-7680

and their tag clouds are depicted in Figure 1. Tag clouds display the most popular tags. The bigger the font size of a tag, the more learning resources are indexed with the tag.

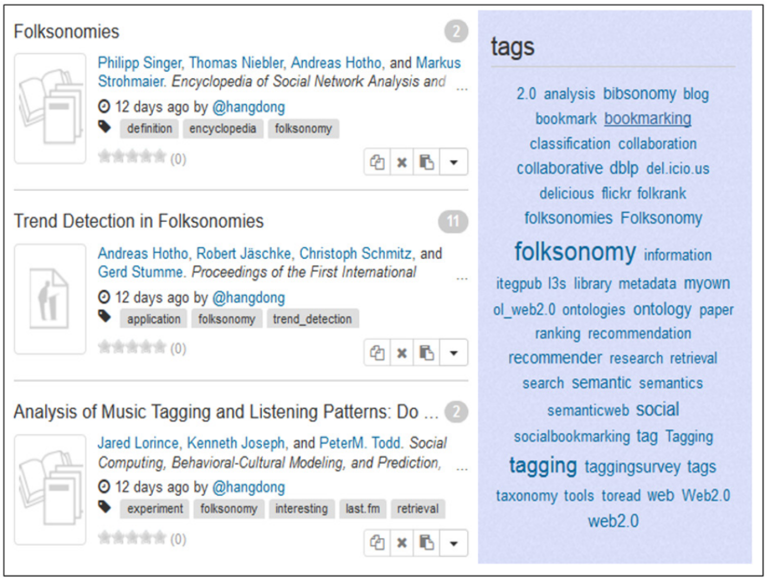

Figure 1. Example of Learning Resources Tagged by Learners and Their Tag Cloud

Since learners personally assign tags to the learning resources, the tags are often ambiguous, overly personalised and imprecise. In addition, learners may use acronyms words, conjugated words, compound words and nonsense words to tag their learning resources and these tags may be shared among a group of learners in collaborative learning environments. This has resulted in uncontrolled and chaotic sets of tagging terms (uncontrolled vocabularies) that cause learning resources retrieval, reuse and sharing among learners to become ineffective and inefficient (Macgregor & McCulloch, 2006; Koren, 2010; Gueye et al., 2014;). Hence, the use of controlled vocabularies to tag learning resources should be prominent prior to knowledge sharing and reuse in a collaborative learning environment (Lau et al., 2015).

In a social tagging system, there are three typical approaches of tag recommendation for annotating learning resources as listed below:

1.) Content-based tagging approach – recommends a tag to a user based on items with similar content in the user's profile. Relies on the content of the documents (Figure 2(a) & Figure 2(b)).

2.) Collaborative-based tagging approach – recommends a tag to a user based on tags used by similar users. Relies on the tagging behaviour of similar users (Figure 2(c)).

3.) Hybrid or combined approaches.

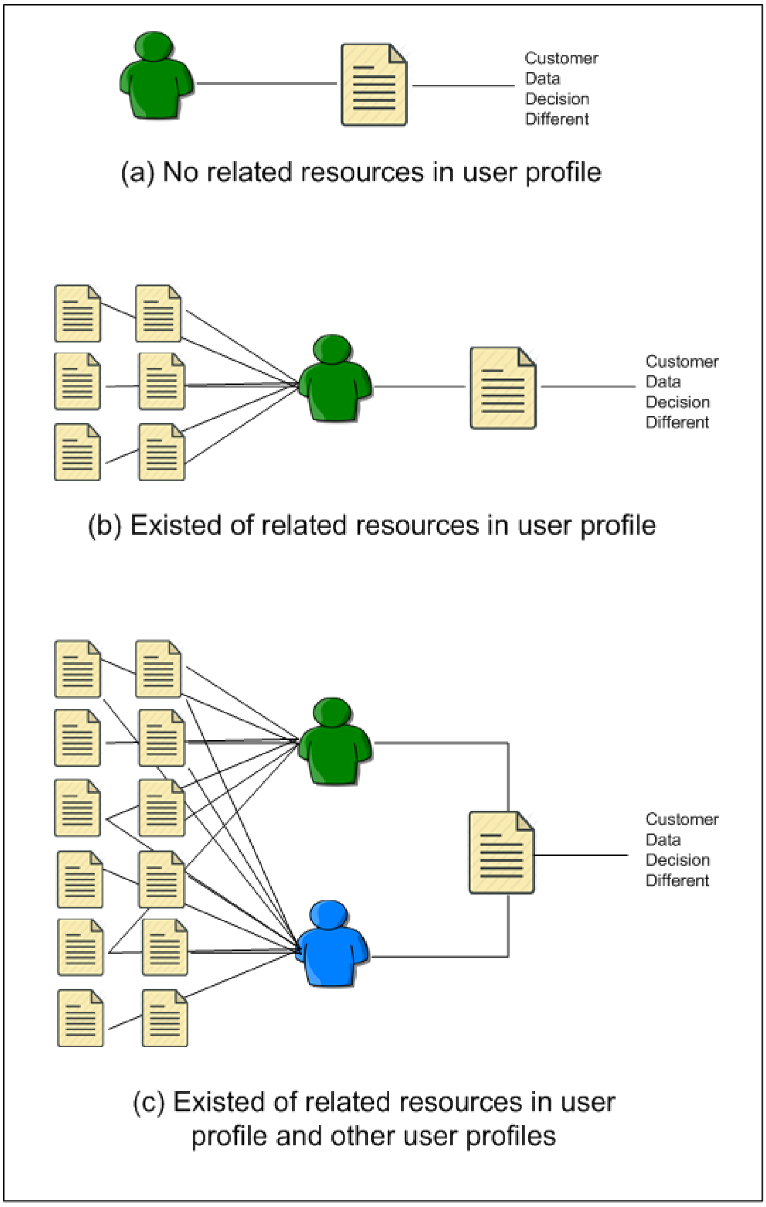

Figure 2. Tag Recommendation Approaches of Tagging System

Automatic tag recommendations can ease the task of learners when annotating new resources. The majority of the tag recommendation approaches assume that a learning resource posting by a learner already exists in the system (prior knowledge) (Jaschke et al., 2007) as depicted in Figure 2(b) and Figure 2(c). In the case where there is no prior knowledge and a learner posts for the first time, the tag recommendation needs to rely on the content of a learning resource to provide good recommendations for unseen resources (Lipczak, 2008; Lu et al., 2009; Tatu et al., 2008; Hassan et al., 2009) as shown in Figure 2(a).

This paper addresses the limitation of uncontrolled vocabularies and the absence of prior knowledge in tagging a learning resource by presenting a content-based approach to automatically generate social tags (controlled vocabularies) from a learning resource using Part-Of-Speech Tagging and K-Means Clustering techniques. The generated social tags from a learning content itself are regarded as controlled vocabularies, hence they are more precise and unambiguous social tags to use for tagging learning resources. The searching, reuse and sharing on the learning resources among learners becomes more effective and efficient through the generated social tags.

The paper is organized as follows: In next section, we discuss related work. Following section explains methodologies for tag generation which comprise folksonomy, part-of-speech tagging and k-means clustering. The framework for automatic tag generation is presented in next section. An example of the results of a simulation to illustrate the framework algorithm is provided in following section. Lastly, the paper is concluded with conclusions and future work.

RELATED WORK

Automatic tag recommendations can reduce people's tagging effort and encourage them to use more tags to annotate resources in a folksonomy system. Therefore, the annotated learning resources can be easily retrieved, shared and reused among people in the folksonomy system. Tags provide new information to resources over original contents (Bischoff et al., 2008; Zhang & Ge, 2015; De Caro et al., 2016), hence tags enhance the capability to discover relevant resources via existing search engines (Heymann et al., 2008; Wei et al., 2016). Furthermore, automatic tag recommendations lighten the task of users while annotating a new resource when prior tag information is not available.

Most of the tag recommendations rely on prior tag information (Sigurbjornsson & Zwol, 2008; Jaschke et. al., 2007; Symeonidis et al., 2008). However, if prior tag information is not available, then the contents of the posted resource need to be relied upon. For folksonomy systems, contents of resources are textual and unstructured in nature, hence appropriate text and natural language processing techniques are required to overcome them (Hassan et. al., 2009). In addition, clustering techniques have been used to handle parsing of document spaces in summarizing large set of documents (Kogan et al., 2006; Kiu & Eric, 2011) for information retrieval and post recommendation in collaborative tagging systems (Begelman et al., 2006; Shepitsen et al., 2008; Lau et. al., 2015).

As discussed in previous section, social tagging systems mainly can be categorized into two approaches, the content-based tagging approach and collaborative-based tagging approach. In this section, content-based tagging for folksonomy systems proposed by Lipczak (2008), Tatu et. al. (2008), Heymann et. al. (2008), Hassan et. al. (2009) and Lu et. al. (2009) are discussed.

Lipczak (2008) proposed a method to extract the terms in the title of a post, and then to expand the set using a tag co-occurrence database. The result is filtered with the poster's tagging history. Meanwhile, Tatu et. al. (2008) used terms from several fields including URL and title to build post and user based models. Natural language processing is used to normalize terms from various sets before recommending them.

Heymann et. al. (2008) developed content-based tag recommendation using a supervised learning method. Page text, anchor text, surrounding hosts and available tag information are formulated as training data. They trained a classifier for each tag they wanted to predict. However, the time required to train the classifiers for each tag becomes substantial when the number of distinct tags increases. For little tag information associated with documents, an association rules mining is used to generate the tag set of the document.

Hassan et. al. (2009) implemented a discriminative clustering approach for content-based tag recommendation in social bookmarking systems. They grouped posts based on the textual contents of the posts with discriminative clustering. The clustering method was used to build two clustering models whereby the first cluster was based on the tags assigned to posts and the second cluster was based on the content terms of posts. In the case of a new posting, the clustering method generated a ranked list of tags which served as final tag recommendations. If the tagging history is available, then this list is also utilized in the final tag recommendation.

Lu et. al. (2009) proposed a content-based tag recommendation for users to manually and automatically annotate webpages with or without prior tag information. Each webpage shares the tags they own with similar webpages. The similarity metric between sending and receiving webpages is defined as a linear combination of four cosine similarities. The similarity is calculated based on tag information and page content. An entropy-based metric is used to describe tags/terms to represent the annotated document. They represent each document with two vectors, namely a tag vector and a term vector using a vector space model.

There are also many content-based tagging systems with prior tag information consisting of different methods (Zhang et al., 2009; Guan et al., 2009; Wetzker et al., 2010; Rendle & Schmidt-Thieme, 2010). Collaborative tagging approaches are seen in studies by Chen et al. (2008), Koren (2010), Gueye et al. (2014) and Ifada, & Nayak (2015).

We present a content-based approach which automatically generates social tags from a learning resource using Part-Of-Speech Tagging and K-Means Clustering techniques to tag learning resources without prior tag information available in social tagging system.

METHODOLOGIES FOR SOCIAL TAG GENERATION

In this section, we provide a formal description of folksonomy, part-of-speech tagging and k–means clustering algorithm.

Folksonomy

A folksonomy is a system where users can use personal or public tags to annotate online resources such as web pages, videos, podcasts, photos and others. It is also known as collaborative tagging and social tagging. This social tagging system uses these tags to index information, facilitate searches and navigate resources (Wikipedia, 2015).

As depicted in Figure 3, a folksonomy consists of users, tags, documents (resources) and the user-based assignment of tags to resources. A folksonomy can be written as a tuple $F := (U, T, R, Y)$ where U (users), T (tags), and R (documents) are finite sets, and Y is a ternary relation between the three components, $Y \subseteq U \times T \times R$, (Jäschke et al., 2007).

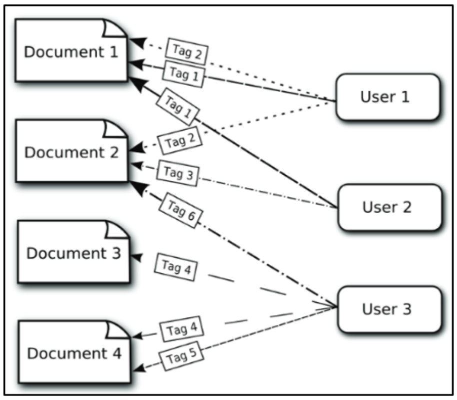

Figure 3. Relations of Tags, Users and Resources (Peters & Stock, 2007)

Part-of-Speech Tagging

Part-of-speech (POS) tagging is applied to transform unstructured text format into a structured text format (Navigli, 2009). Every word in the text is assigned a unique part-of-speech as listed in Table 1. It predicts the part-of-speech even for an unknown word by exploiting the context of the word in a sentence. The POS tagging process is shown in Figure 4 and explicated as below:

1.) Tokenization is a normalization process that splits up the text into a set of words.

2.) Part-of-speech tagging is the process to assign a grammatical category to each word. This process is referred as grammatical tagging as it assigns a single part-of-speech tag to each word and punctuation marker.

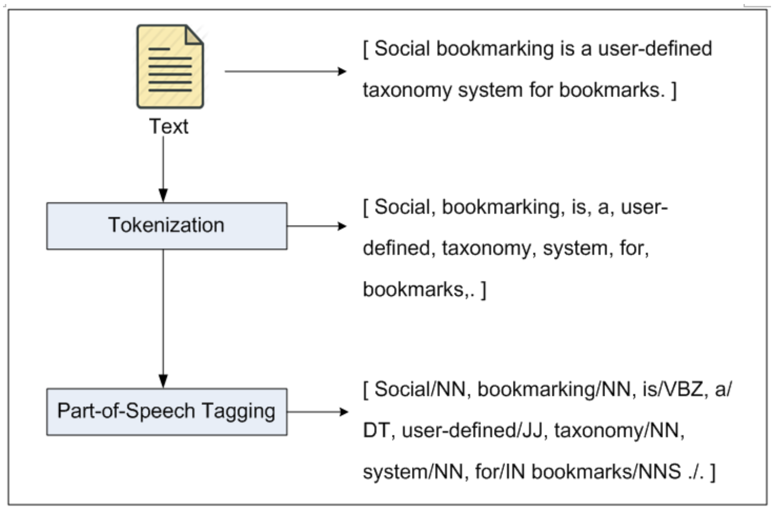

Figure 4. Process of POS Tagger

Table 1
Description for Tags of Part-of-Speech

Tag	Description	Tag	Description
$	dollar	PDT	pre-determiner
``	opening quotation mark	POS	genitive marker
"	closing quotation mark	PRP	pronoun, personal
(	opening parenthesis	PRP$	pronoun, possessive
)	closing parenthesis	RB	adverb
,	comma	RBR	adverb, comparative
--	dash	RBS	adverb, superlative
.	sentence terminator	RP	particle
:	colon or ellipsis	SYM	symbol
CC	conjunction, coordinating	TO	"to" as preposition or infinitive marker
CD	numeral, cardinal	UH	interjection
DT	determiner	VB	verb, base form
EX	existential there	VBD	verb, past tense
FW	foreign word	VBG	verb, present participle or gerund
IN	preposition or conjunction, subordinating	VBN	verb, past participle
JJ	adjective or numeral, ordinal	VBP	verb, present tense, not 3rd person singular
JJR	adjective, comparative	VBZ	verb, present tense, 3rd person singular
JJS	adjective, superlative	WDT	WH-determiner
LS	list item marker	WP	WH-pronoun
MD	modal auxiliary	WP$	WH-pronoun, possessive
NN	noun, common, singular or mass	WRB	Wh-adverb
NNP	noun, proper, singular		
NNPS	noun, proper, plural		
NNS	noun, common, plural		

K-Means Clustering

K-Means clustering is a simple unsupervised clustering technique (Berkhin, 2006). It is used to classify a given data set into a previously specified number of clusters (*k*). K-means clustering divides a data set to a number of clusters. It is defined as

$$E = \sum_{k=1}^{C} \sum_{x \in Q_k} \|x - c_k\|$$

where C is the number of clusters, x is a data point, and c_k is the centroid of the data points k.

The algorithm of k-means is composed of four steps:

Step 1: Specify the number of clusters, *k*. Initialize *k* point randomly as cluster centers.
Step 2: Assign each instance to its closest cluster center using Euclidean distance.

$$Euclidean\ distance: dist(x, \mu) = \sum_{j=i}^{d} (x_j - \mu_j)^2$$

where μ_j is mean of point in x_j.

Step 3: Re-compute the centroid (mean) for each cluster as a new cluster center.

$$mean: \vec{\mu} = \frac{1}{N_C} \sum_{i \in C} \vec{X_i}$$

Step 4: If the new cluster centers are different from the old cluster centers repeat Step 2 until the cluster centers do not change anymore.

THE SOCIAL TAG GENERATION FRAMEWORK

This section explicates the automatic social tag generation framework as depicted in Figure 5.

The algorithm implementation of the framework comprises the following steps:

Input : A learning resource

Step 1 : HTML Parser
A learning resource obtained online is parsed into unstructured text (unformatted text) using HTML Content Extractor (Alexander, 2015) automatically. All non-text contents such as flash animation, hyperlinks, audios, videos and others are parsed. Meanwhile the unformatted text context is saved into text processing format (.txt) as output for the next process.

Step 2 : Tokenization
The sentences are tokenized into each word (token) prior to the POS tagging process. For example, the text "`This is social bookmarking.`", is tokenized into "`This, is, a, social, bookmarking, .`".

Step 3 : Part-of-Speech Tagging
The tokenized words are assigned with POS tags. The Standford POS English tagger, namely bidirectional-distsim (Toutanova et. al, 2003) is used to perform the POS tagging assignment. For example, the tokens "`This, is, a, social, bookmarking, .`", are POS tagged into "`This_DT is_VBZ a_DT Social_NNP bookmarking_NN ._.`". The structured text is generalized at the end of this process.

Step 4 : Contextual Clustering
K-means clustering in Weka is used to contextualize the structured text in order to reduce the tag generation processing time (Hall et. al., 2009).

Clusters of five are applied in k-means clustering to cluster each word in the structured text. The words are clustered into five clusters according to their attributes.

Step 5 : Keyword Generation
In this process, words that are associated with the unconventional part-of-speech (`` , :, ? CD, DT, VBZ, IN, PRP, RBS, VBN, VBP, EX, -RRB-, -LRB-, CC, VBG, MD, VB, WRB, RB, WDT, RP, VBD, RP, WP, JJR, TO, PRP$, POS, RP) in each cluster are eliminated to retain only the words in nouns. Key words (tags) are formed based on the sequence relations of each word in a cluster.

Output : A set of tags (controlled vocabularies).

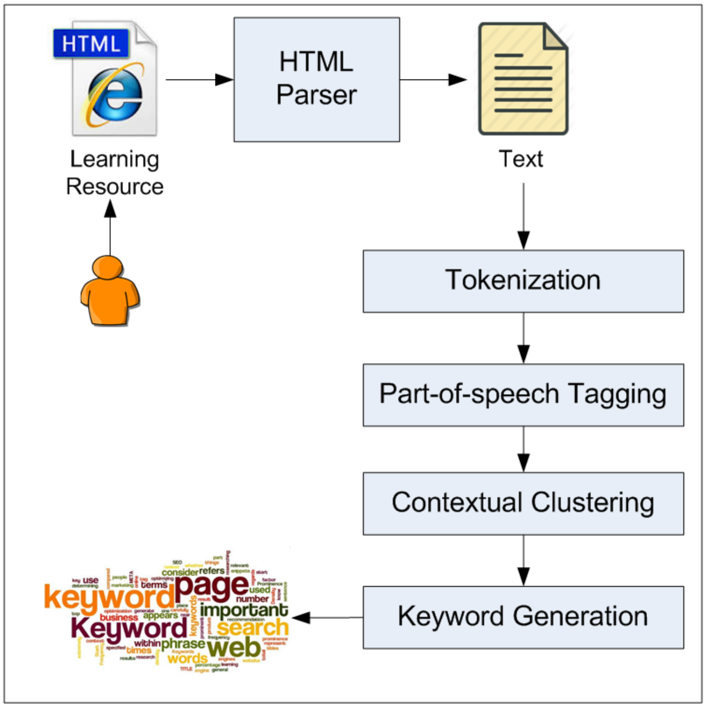

Figure 5. Framework for Automatic Social Tag Generation

SIMULATED EXAMPLE

The learning resource (http://whatis.techtarget.com/definition/social-bookmarking) to be tagged is illustrated in Figure 6. The learning resource is parsed into unstructured text using HTML Parser and saved into text processing format. Each of the words in the file is tokenized and assigned with part-of-speech has resulted 341 tokens as illustrated in Figure 7. K-Mean clustering (Figure 8) is applied to cluster the tokens according to the tokens' attributes. 117 token are yielded after the process of contextual clustering (Figure 9). The process of elimination the unconventional part-of-speech which associated with the words has generated 23 tags (controlled vocabularies) as shown in Figure 10. The generated tags are the recommended tags for a learner to annotate the learning resource.


DEFINITION

social bookmarking

Like 3 Tweet 8+1 0 Share Email

Part of the Internet applications glossary:

Social bookmarking is a user-defined taxonomy system for bookmark s. Such a taxonomy is sometimes called a folksonomy and the bookmarks are referred to as tags. Unlike storing bookmarks in a folder on your computer, tagged pages are stored on the Web and can be accessed from any computer. Technorati, a blogging site, describes the system as "The real-time Web, organized by you." Web sites dedicated to social bookmarking, such as Flickr and del.icio.us, provide users with a place to store, categorize, annotate and share favorite Web pages and files.

Here's how you might use del.icio.us: Once you've registered, you can install buttons to your toolbar . If you visit a Web page that you want to save, you click a "tag" button. A window pops up with the URL and page name and optional fields for you to add notes and tags. Tags are single-word descriptors -- this is how social bookmark systems categorize content so that you and other users can find it.

Another button takes you directly to a page with your saved bookmarks and associated tags. If you click on one of the tags you created, you'll be taken to a page listing your saved pages associated with that tag. You can also view all users' content associated with that tag or the most popular examples. The site inbox feature lets you subscribe to content saved with a specific tag or content saved by a specific user or a combination of the two.

When you click a tag or search for a term on del.icio.us, the results include information about how many members have saved each item and any tags that they included for it.

Joshua Schachter started del.icio.us in 2003 as a way to track and share bookmarks. As the service grew and demanded more of his time, Schachter quit his job at Morgan Stanley to devote himself to the site, which now serves hundreds of thousands of members. In December of 2005, he sold the site to Yahoo.

This was last updated in June 2006

Posted by: Margaret Rouse

Figure 6. The Tagged Learning Resource

Index	Token	Tag	Index	Token	Tag	Index	Token	Tag
1	Social	NNP	21	bookmark	NNS	41	the	DT
2	bookmarking	NN	22	are	VBP	42	Web	NN
3	is	VBZ	23	referred	VBN	43	and	CC
4	a	DT	24	to	TO	44	can	MD
5	user-defined	JJ	25	as	IN	45	be	VB
6	taxonomy	NN	26	tags	NNS	46	accessed	VBN
7	system	NN	27	Unlike	IN	47	from	IN
8	for	IN	28	storing	VBG	48	any	DT
9	bookmark	NN	29	bookmark	NNS	49	computer	NN
10	s	NN	30	in	IN	50	Technorat	NNP
11	Such	JJ	31	a	DT	:	:	:
12	a	DT	32	folder	NN	:	:	:
13	taxonomy	NN	33	on	IN	:	:	:
14	is	VBZ	34	your	PRP$	:	:	:
15	sometimes	RB	35	computer	NN	336	he	PRP
16	called	VBN	36	tagged	VBN	337	sold	VBD
17	a	DT	37	pages	NNS	338	the	DT
18	folksonomy	NN	38	are	VBP	339	site	NN
19	and	CC	39	stored	VBN	340	to	TO
20	the	DT	40	on	IN	341	Yahoo	NNP

Figure 7. List of Tokens after POS Tagger

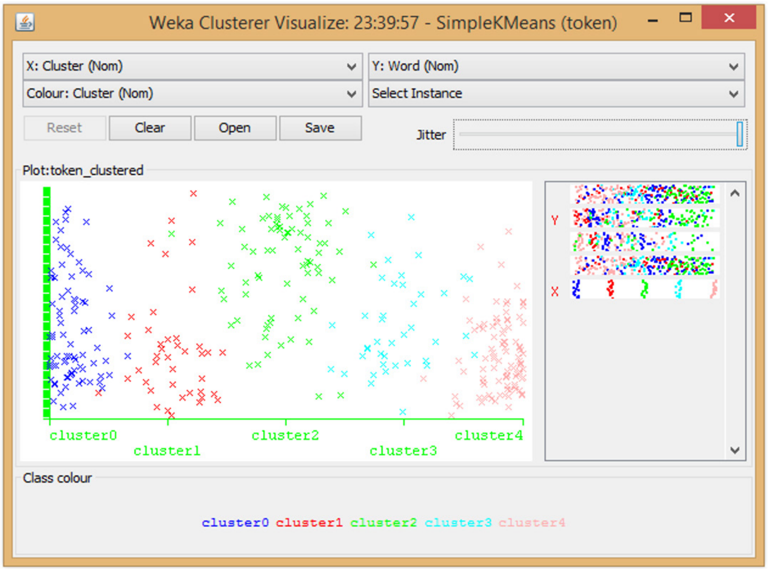

Figure 8. Visualization of Clustered Tokens

Index	Token	Tag	Cluster	Index	Token	Tag	Cluster	Index	Token	Tag	Cluster
12	Social	NNP	cluster1	21	real-time	JJ	cluster1	:	:	:	:
13	bookmarking	NN	cluster1	22	Web	NN	cluster1	:	:	:	:
14	user-defined	JJ	cluster1	23	Web	NN	cluster1	:	:	:	:
15	taxonomy	NN	cluster1	24	sites	NNS	cluster1	:	:	:	:
16	system	NN	cluster1	25	social	JJ	cluster1	102	track	NN	cluster1
17	bookmark	NN	cluster1	26	bookmark	NN	cluster1	103	share	NN	cluster1
18	Such	JJ	cluster1	27	such	JJ	cluster1	104	bookmark	NNS	cluster4
19	taxonomy	NN	cluster1	28	Flickr	NNP	cluster1	105	service	NN	cluster1
20	folksonomy	NN	cluster1	29	del.icio.us	NNP	cluster1	106	time	NN	cluster1
22	tags	NNS	cluster1	30	users	NNS	cluster1	107	Schachter	NNP	cluster4
23	bookmarks	NNS	cluster1	31	place	NN	cluster1	108	job	NN	cluster1
24	folder	NN	cluster1	32	store	NN	cluster1	109	Morgan	NNP	cluster4
25	computer	NN	cluster1	33	categorize	NN	cluster1	110	Stanley	NNP	cluster4
26	pages	NNS	cluster1	34	annotate	NN	cluster1	111	site	NN	cluster1
27	Web	NN	cluster1	35	share	NN	cluster1	112	hundreds	NNS	cluster4
28	computer	NN	cluster1	36	favorite	JJ	cluster1	113	thousands	NNS	cluster4
29	Technorati	NNP	cluster1	37	Web	NN	cluster1	114	members	NNS	cluster4
30	blogging	NN	cluster1	38	pages	NNS	cluster1	115	December	NNP	cluster4
31	site	NN	cluster1	39	files	NNS	cluster1	116	site	NN	cluster1
32	system	NN	cluster1	40	del.icio.us	NN	cluster1	117	Yahoo	NNP	cluster4

Figure 9. List of Tokens after Contextual Clustering and Part-of-Speech Elimination

Tags (Controlled Vacabularies)		
blogging site	Morgan Stanley	site inbox feature
bookmarks	optinal fields	social bookmarking
buttons	other users	systems
computer	page name	tags
content	pages	user-defined taxonomy system
del.icio.us	real-time web	web pages
Joshua Schachter	share favorite web pages	web sites
members	single-word descriptors	

Figure 10. Generated Tags for a Learning Resource Tagging

CONCLUSION AND FUTURE WORK

Tags used by learners to annotate learning resources are often ambiguous, imprecise and overly personalised. Hence this has resulted in uncontrolled vocabularies that cause learning resources retrieval, sharing and reuse to become ineffective. This paper has proposed an automatic tags generation algorithm to generate controlled vocabularies (social tags) from contents of a learning resource. The experimental result has shown the used of part-of-speech and unsupervised clustering techniques in the proposed algorithm has improved the scalability of tag generation from a learning resource. Furthermore, prior knowledge (tag information) is not required in tag generation. For future work, we will investigate a suitable weighting scheme for tag recommendation by incorporating a learner profile.

Tagging a learning resource with the generated controlled vocabularies can appropriately describe the content of the learning resource itself. Subsequently, these learning resources can be shared and reused effectively and efficiently among learners in an e-learning environment. In addition, learning resources navigation and retrieval are more effective with controlled vocabularies annotation.

The algorithm can be applied to generate tags for explicit knowledge organizing and navigating in knowledge management systems. It also can be extended to E-Commerce system to generate tags from product catalogue or description. Product tagging with controlled vocabularies can enhance product searching, recommendation and personalization. On the other hand, the generated tags from a product catalogue can serve as SEO keywords and metadata for the product itself.

REFERENCES

Alexander, B. (2015). *HTML-Content-Extractor*. Retrieved 1 November 2015, from http://search.cpan.org/dist/HTML-Content-Extractor/.

Begelman, G., Keller, P., & Smadja, F. (2006). Automated tag clustering: Improving search and exploration in the tag space. In *Collaborative Web Tagging Workshop at WWW2006.* Edinburgh, Scotland, UK.

Berkhin, P. (2006). A survey of clustering data mining techniques. In Kogan, J., Nicholas, C., & Teboulle, M. (Eds.). *Grouping multidimensional data*. (25-71). Berlin Heidelberg: Springer.

Bischoff, K., Firan, C. S., Nejdl, W., & Paiu, R. (2008). Can all tags be used for search? In *Proceedings of the 17th ACM conference on Information and knowledge management* (193-202). New York: ACM.

Chen, W.Y., Zhang, D., & Chang, E.Y. (2008) Combinational Collaborative Filtering for Personalized Community Recommendation. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Las Vegas, 24-27 August 2008* (115-123). New York: ACM. http://dx.doi.org/10.1145/1401890.1401909

De Caro, P., Pini, M. S., & Sambo, F. (2016). A Tag-Based Recommender System. In Menegatti, E., Michael, N., Berns, K., & Yamaguchi, H. (Eds). *Intelligent Autonomous Systems 13* (1049-1061). Cham, Switzerland: Springer International Publishing.

Guan, Z., Bu, J., Mei, Q., Chen, C., & Wang, C. (2009) Personalized Tag Recommendation Using Graph-Based Ranking on Multi-Type Interrelated Objects. In *Proceedings of the 32nd International*

ACM SIGIR Conference on Research and Development in Information Retrieval, New York, 19-23 July 2009 (540-547). New York: ACM. http://dx.doi.org/10.1145/1571941.1572034

Gueye, M., Abdessalem, T., & Naacke, H. (2014). A Parameter-Free Algorithm for an Optimized Tag Recommendation List Size. In *Proceedings of the 8th ACM Conference on Recommender Systems* (233-240). New York: ACM.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter*, *11*(1), 10-18.

Hassan, M. T., Karim, A., Manandhar, S., & Cussens, J. (2009). Discriminative clustering for content-based tag recommendation in social bookmarking systems. *ECML PKDD Discovery Challenge 2009 (DC09)* (85). Retrieved 1 November, 2015, from http://ceur-ws.org/Vol-497

Heymann, P., Koutrika, G., & Garcia-Molina, H. (2008). Can social bookmarking improve web search? In *Proceedings of the 2008 International Conference on Web Search and Data Mining* (195-206). New York: ACM.

Ifada, N. & Nayak, R. (2015) Do-Rank: DCG Optimization for Learning-to-Rank in Tag-Based Item Recommendation Systems. In: Cao, T., Lim, E.-P., Zhou, Z.-H., Ho, T.-B., Cheung, D. & Motoda, H., Eds., *Advances in Knowledge Discovery and Data Mining* (510-521). Berlin Heidelberg: Springer International Publishing. http://dx.doi.org/10.1007/978-3-319-18032-8_40

Jaschke, R., Marinho, L., Hotho, A., Schmidt-Thieme, L., & Stumme, G. (2007). Tag recommendations in folksonomies. In *Knowledge Discovery in Databases: PKDD 2007* (506-514). Berlin Heidelberg: Springer.

Kiu, C. C., & Tsui, E. (2011). TaxoFolk: A hybrid taxonomy–folksonomy structure for knowledge classification and navigation. *Expert Systems with Applications*, *38*(5), 6049-6058.

Kogan, J., Nicholas, C., & Teboulle, M. (Eds.) (2006). *Grouping multidimensional data*. Berlin Heidelberg: Springer.

Koren, Y. (2010) Collaborative Filtering with Temporal Dynamics. *Communications of the ACM* , *53*, 89-97. http://dx.doi.org/10.1145/1721654.1721677

Lau, S. B. Y., Lee, C. S., & Singh, Y. P. (2015). A folksonomy-based lightweight resource annotation metadata schema for personalized hypermedia learning resource delivery. *Interactive Learning Environments*, *23*(1), 79-105.

Lipczak, M. (2008). Tag recommendation for folksonomies oriented towards individual users. In *ECML PKDD Discovery Challenge 2008* (84–95). Retrieved 1, November 2015, from http://www.kde.cs.uni-kassel.de/ws/rsdc08

Lu, Y. T., Yu, S. I., Chang, T. C., & Hsu, J. Y. J. (2009). A Content-Based Method to Enhance Tag Recommendation. *IJCAI 9*, 2064-2069.

Macgregor, G., & McCulloch, E. (2006). Collaborative tagging as a knowledge organisation and resource discovery tool. *Library review*, *55*(5), 291-300.

Navigli, R. (2009). Word sense disambiguation: A survey. *ACM Computing Surveys (CSUR)*, *41*(2), 10.

Peters, I., & Stock, W. G. (2007). Folksonomy and information retrieval. In *Proceedings of the 70th Annual Meeting of the American Society for Information Science and Technology*, *45*, 1510-1542.

Rendle, S. & Schmidt-Thieme, L. (2010) Pairwise Interaction Tensor Factorization for Personalized Tag Recommendation. In *Proceedings of the 3rd ACM International Conference on Web Search and Data Mining, 2010* (81-90). New York: ACM. http://dx.doi.org/10.1145/1718487.1718498

Shepitsen, A., Gemmell, J., Mobasher, B., & Burke, R. (2008). Personalized recommendation in social tagging systems using hierarchical clustering. In *Proceedings of the 2008 ACM conference on Recommender systems* (259-266). New York: ACM.

Sigurbjornsson, B. & Van Zwol, R. (2008). Flickr tag recommendation based on collective knowledge. In *Proceedings of the 17th international conference on World Wide Web*. (327-336). New York: ACM.

Symeonidis, P., Nanopoulos, A., & Manolopoulos, Y. (2008). Tag recommendations based on tensor dimensionality reduction. In *Proceedings of the 2008 ACM conference on Recommender systems* (43-50). New York: ACM.

Tatu, M., Srikanth, M., & D'Silva, T. (2008). Rsdc'08: Tag recommendations using bookmark content. *ECML PKDD discovery challenge*, 96-107.

Toutanova, K., Klein, D., Manning, C. D., & Singer, Y. (2003). Feature-rich part-of-speech tagging with a cyclic dependency network. In *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1* (173-180) Stroudsburg PA: Association for Computational Linguistics.

Wei, S., Zheng, X., Chen, D., & Chen, C. (2016). A hybrid approach for movie recommendation via tags and ratings. *Electronic Commerce Research and Applications*. doi: 10.1016/j.elerap.2016.01.003

Wetzker, R., Zimmermann, C., Bauckhage, C., & Albayrak, S. (2010). I Tag, You Tag: Translating Tags for Advanced User Models. In *Proceedings of the 3rd ACM International Conference on Web Search and Data Mining, New York, 2010* (71-80). New York: ACM. http://dx.doi.org/10.1145/1718487.1718497

Wikipedia (2015). Folksonomy. Retrieved 25 March 2016, from https://en.wikipedia.org/wiki/Folksonomy#cite_note-Berlin.2C_B._1992-21

Zhang, N., Zhang, Y., & Tang, J. (2009). A Tag Recommendation System for Folksonomy. In *Proceedings of the 2nd ACM Workshop on Social Web Search and Mining, New York, 2 November 2009* (9-16). New York: ACM.

Zhang, S. & Ge, Y. (2015). Personalized Tag Recommendation Based on Transfer Matrix and Collaborative Filtering. *Journal of Computer and Communications*, *3*(09), 9.