

Cultural Difference on the Usage Pattern of Tagging System for Knowledge Sharing

Zhe Chen, Qin Gao, and Yuancheng Yang

Institute of Human Factors & Ergonomics, Industrial Engineering,
Tsinghua University, China
{chenzhe.demi, yangyuancheng}@gmail.com,
gaoqin@tsinghua.edu.cn,

Abstract. Collaborative tagging system in online knowledge sharing system benefits the knowledge management of users. It also inspires us a quantitative way how to investigate the effect of culture on knowledge management as tags reflects the usage pattern of users. In this study, the usage pattern of tag in three types of knowledge management websites (i.e. photo, bookmark and book sharing websites) present the influence of Chinese culture upon knowledge management. A set of indices is used to describe the usage pattern of tags and the results indicate that Chinese users behave differently in all three types of knowledge sharing system.

Keywords: tag, knowledge sharing, Chinese culture.

1 Introduction

Knowledge management has recently gained its prevalence on the web due to collaborative tagging system. Users attach tags to almost any content in the Internet and share these tags to others. A new information categorization method based on tagging is created. This method provides a convenient, fast, personal way for information's retrieval, filtering, and navigation, instead of traditional taxonomy.

The collective tagging system also fosters a quantitative way in the research of knowledge management. Many studies inspired by tagging systems have been conducted. However, few studies considered the effect of culture. This study investigates the cultural effect on tag usage pattern of Chinese user.

Globalization makes communication and co-work between different cultures more common [6]. Cross-culture team is everywhere. So it is necessary to know how culture affects knowledge management activities of user [7]. Collective tagging provides a new perspective to analyze the effect of culture.

2 Literature Review

Sharing of knowledge as knowledge defined as “fluid mix of framed experience, values, contextual information, and expert insight that provides a framework for

evaluating and incorporating new experiences and information" [4], are a hot spot in the field of knowledge management. Knowledge sharing on web usually indicates the contribution of knowledge and meanwhile acquirement of others contribution. There are many works at deriving an effective way of knowledge management.

As more and more web 2.0 applications spring up, an online approach for knowledge management and sharing emerges. Collective tagging provides us a new way to organize information, especial categorization. On the contrast of taxonomy, collective tagging of categorization (folksonomy) is considered to be a radical and more effective way of classify contents or documents web [15]. The lack of authority of professional probably leads to chaos of information, rather than form regularities [10].

Tag on the web means an unrestrained descriptor or keyword, added to item, in forms of alphanumeric and symbols, while tagging implicates user's behavior of adding tags. The most distinguishing characteristic of the tag is that it is freely-formed as users use their own terms. Tags are then used for future navigation, filtering and retrieval, which provide a new method of information organization. Tagging for knowledge management becomes prevalent, as the encouragement of Web 2.0 applications. For example, in the del.icio.us (a tagging system), user bookmarked websites with their own descriptors. Thus, users create their personal categorization. Then the stable usage pattern has emerged from collective tagging [5].

Del.icio.us, as a popular website of collective tagging system, is for users to organize their favorite bookmarks online, using description and tag. An empirical research on Del.icio.us indicates that there are stable patterns of tag proportion [5]. Another empirical research on Del.icio.us uses multi-dimensional scaling to analyze tag frequency and co-word pattern of collective tagging [10]. Cameron et al offer us a conceptual model of tagging system, assisting for analysis in depth [13].

As culture has the effect on individual behavior, so it is of little possibility culture has negligible effect on the regularities of users' knowledge sharing activities. Then it is interesting to investigate the effect of culture in the knowledge sharing system [9].

Chinese culture is distinctly different from other cultures. Chinese culture vitiate greatly from other culture and Chinese tend to be more collective than people from western cultures [8]. Chiu [2] presents the differences cognitive styles between Chinese and American, and it refers to that Americans are Inferential-categorical while Chinese are relational-contextual. Chinese are more used to living on a social-relational organization than people from western cultures. Yang [16] describes personality of Chinese as restrained, patient, and self-contained and American as more impulsive, excitable and spontaneous. Moreover, the holistic perspective into objects gives Chinese a wider scope of the world [1]. It is found that Chinese users contribute less than users from other cultures in knowledge sharing websites [3]. Chinese users are likely to show a different pattern on different types of knowledge sharing websites [1].

The objective of this study is to investigate the effect of Chinese culture on knowledge sharing system through the usage pattern of tagging system. The results of this study can be used to improve knowledge sharing the cross-cultural team in global organizations.

3 Method

Marlow et al. state that tag system has three key elements (i.e. Resources, Tag and User) [13]. Then following indices are used to describe the tag usage in tag system as shown in table 1.

Table 1. The indices of usage pattern in tagging system

Abbr.	Meaning of Index
T	Number of Unique tags
UU	Number of Uploading User
UT	Number of tagging user
U	Number of user - user of tag system, including uploading user and tagging user
I	Number of unique item - given to some user share resources without tagging them, some items have no tags attached to them.
NT	Number of tag been used - It record 1 time as a user tag a item using a certain tag.
NU	Number of user tagging - It record 1 time as a certain user tag a item using a tag.
NI	Number of item been tagged - It record 1 time as a user tag a certain item using a tag.
T_U	Number of unique tags for a certain user
I_U1	Number of unique items for a certain user
I_U2	Number of unique tagged items for a certain user
T_I	Number of unique tags for a certain item
U_I1	Number of unique users for a certain item
U_I2	Number of unique tagging users for a certain item
U_T	Number of unique users for a certain tag
I_T	Number of unique items for a certain tag
T_IU	Number of unique tags for a certain item of a certain user
U_TI	Number of unique users for a certain tag of a certain item
I_TU	Number of unique items for a certain tag of a certain user
RU1	Level of repetition that computed as number of unique tags for a certain item of a certain user divided by number of unique items for that user
RU2	Level of repetition that computed as number of unique tags for a certain item of a certain user divided by number of unique tagged items for that user
RI1	Level of repetition that computed as number of unique users for a certain tag of a certain item divided by number of unique users for that item
RI2	Level of repetition that computed as number of unique users for a certain tag of a certain item divided by number of unique tagging users for that item

In the collective knowledge-sharing tagging system, items such as book, websites, photo, and video can be shared with other users. Since both Chinese websites and English websites should have similar tagging systems and users' data can be acquired from webpage, three kinds of knowledge sharing websites are selected for comparison of usage pattern of tag (i.e. photo, website, book sharing websites). Introduction to these websites is in Table 2.

Table 2. Introduction to websites for comparison

Website	Categoriza- tion	Data collection	Url	Function
Yupoo	Photo	From website	www.yupoo.com	Uploading, sharing and tagging photos
Flickr	Photo	From website	www.flickr.com	Uploading, sharing and tagging photos
QQ	Website	From website	shuqian.qq.com	Collecting, sharing and tagging websites
Del.icio.us	Website	From website	www.delicious.com	Collecting, sharing and tagging websites
Douban	Book	API	www.douban.com	Collecting, sharing and tagging books
Shelfari	Book	From website	www.shelfari.com	Collecting, sharing and tagging books

Raw data are observed directly from the knowledge sharing website. Cases are recorded in data collection while one case including information about the id of user, the id of the item, and name of tags. In one case, there are only one user, one item and at most one tag as in some case user doesn't attach tags on the item. In photo sharing websites, cases are recorded from recent uploading photos, the most interesting photos of 7 days and 100 random user's all cases. In book sharing websites, cases are recorded from 1000 random user's cases and their 50 books.¹ In bookmark sharing system, cases are recorded from 100 users. The detailed data collection information is shown in table 3.

Table 3. Data collection

	Method	Date
Recent uploading photo	Refresh the recent page for 80 times as there are 28 photos on one page in Yupoo; Refresh the recent page for 85 times as there are 20 photos on one page in Flickr	2010.05
the most interesting photos of 7 days	Refresh the interesting page for 100 times in both Yupoo and Flickr as there are 9 photos on both of them	2010.05
100 random photo user's all cases	Search the key word "a", "b"... "z" and "0", "1"... "9" to get users name and then randomly chose 100 of them	2010.05
100 random photo hot user's all cases	Search the key word "a", "b"... "z" and "0", "1"... "9" to get users name in the most hot user list and then randomly chose 100 of them	2010.05
1000 random book user's all cases (at most 50 books for each user) ²	Search the key word "a", "b"... "z" and "0", "1"... "9" to get users name and then randomly chose 1000 of them	2009.05
100 random bookmark user's all cases	Search the key work "a", "b"... "z" and "0", "1"... "9" to get users name and then randomly chose 100 of them	2009.05

¹ Douban's privacy policy allows that at most 50 books for each user can be observed.

² There are only two tags for 100 random users of Yupoo, hottest users' cases are recorded for better comparison.

4 Results and Discussion

The table 4-5 shows the descriptive result and t-test results of this study.

Most users share less knowledge (photo, book, and bookmark) and fewer users share most knowledge in knowledge sharing websites. As in figure 1-3, it is shown that both Chinese websites users and English websites users have this usage pattern. The curve for this usage pattern is similar to the curve of inverse function. This curve indicates that most users have fewer shared items, tags. Most items have fewer tags. Fewer users contribute most items and tags in knowledge sharing websites.

Chinese users upload more photos in one time. In the recent photo sharing websites, while collecting similar number of photos, there are fewer sharers in Yupoo than in Flickr. The average number of photos shared by users in Yupoo (12.262) is significant form Flickr (2.508) as the p-value for t-test is 0.000. This result implies that Chinese user tend to sharing more photos in one time possibly through fixed devices as PC. User in Flickr is more willing to share their photo into website right after them taking this photo through mobile devices.

Chinese users are more willing to tag all knowledge they share. In Chinese recent uploading photos, the 7 users who have tags tag all the photos they shared in website recently and are willing using the same tag, but tagging users in Flickr don't tag all the photos they share. This demonstrated in the in value of I_T that Yupoo user is significantly higher than Flickr user because some users tag the entire photos they share in Yupoo. In Flickr, there is no such kind of user and they consider tagging as an optional function. This reflects in the value of $R1$ and $R2$. This kind of user will have the same value in $R1$ and $R2$. Table 4-5 show that Chinese user have this kind of behavior when they recently upload photos, share interesting photos of 7 days, share book websites and bookmark sharing websites.

In random users' data, more Chinese users have book sharing behavior, but fewer Chinese users have photo and bookmark sharing behavior. This may be demonstrated in value of U and UT . U and UT of Chinese book user is 579 and 276 respectively, compared to 400 and 39 of Shelfari. This indicates a different usage pattern in different knowledge sharing systems. However, sample of book websites is slightly different from other two random users' samples.

Tagging for other users' shared content shows another unique usage pattern of Chinese users. Photo sharing websites have a unique feature, which all knowledge shared in the websites are created and uploaded by the creator himself/herself. So it is an interesting finding that some users tag other users' photos. In recent uploaded photos and random users' photos, users don't tag other users' photos.

The value U_I2 of recent photos and interesting photos photo in Chinese photo sharing websites is 1. In English websites, only photo recently uploaded have 1 in U_I2 . This indicates that no other user tag the photo except the user who uploads it in Chinese book sharing websites. Exploring user may not check the recent uploaded photo. Indices of $RI1$ and $RI2$ also indicate this kind of usage pattern.

Table 4. The descriptive result and t-test result in photo-sharing websites (Recent, interesting, random user)

Abbr	Recent uploaded (photo)			Interesting of 7 days (photo)			100 random users (photo)		
	Yupoo	Flickr	Test	Yupoo	Flickr	Test	Yupoo	Yupoo *	Flickr
T	21	1179		127	582		0	3466	2779
UU	n						7	100	82
UT	n						0	80	40
U	n						7	105	86
I	n						613	70844	9296
NT	Mean	7.381	t	1.310	1.130	t	-	32.040	9.970
	SD	6.160	p	0.587	0.456	p	-	145.913	78.562
NU	Mean	1.099	t	2.650	5.750	t	0.000	1057.530	322.290
	SD	7.090	p	6.152	8.425	p	-	4177.314	1982.014
NI	Mean	0.090	t	1.720	7.210	t	0.000	1.570	2.980
	SD	0.570	p	2.999	9.003	p	-	2.666	3.864
T_U	Mean	0.149	t	2.140	5.640	t	0.000	40.830	37.140
	SD	0.746	p	4.745	8.308	p	-	117.485	142.894
L_U1	Mean	12.262	t	1.540	1.000	t	87.571	674.850	106.900
	SD	31.844	p	1.229	0.000	p	-	1066.675	555.637
L_U2	Mean	7.286	t	1.619	1.096	t	-	418.232	141.375
	SD	6.130	p	1.203	0.402	p	-	908.728	677.871
T_I	Mean	0.090	t	1.720	7.120	t	0.000	1.570	2.370
	SD	0.570	p	2.999	8.949	p	-	2.666	3.579
U_I1	Mean	1.000	t	1.000	1.350	t	0.011	1.000	1.000
	SD	0.000	p	0.000	1.642	p	-	0.015	0.021
U_I2	Mean	1.000	t	1.000	1.517	t	-	1.000	1.001
	SD	0.000	p	0.000	1.918	p	-	0.012	0.027

Table 4. (continued)

Abbr	Recent uploaded (photo)			Interesting of 7 days (photo)			100 random users (photo)			
	Yupoo	Flickr	Test	Yupoo	Flickr	Test	Yupoo	Yupoo *	Flickr	Test
U_T	Mean	1.000	1.128 t (7.646)	1.060	1.100 t (1.316)		-	1.240	1.150 t	5.185
	SD	0.000	0.575 p 0.000*	0.302	0.412 p 0.189		-	0.803	0.525 p	0.000*
L_T	Mean	7.381	1.964 t 4.025	1.310	1.110 t 3.656		-	32.040	7.920 t	8.934
	SD	6.160	2.276 p 0.001*	0.587	0.436 p 0.000*		-	145.913	56.414 p	0.000*
T_IU	Mean	0.717	1.348 t (6.340)	1.720	5.330 t (4.599)		0.000	1.570	2.980 t (34.228)	
	SD	2.692	3.663 p 0.000*	2.999	8.027 p 0.000*		-	2.666	3.863 p	0.000*
U_TI	Mean	1.000	1.000 t -	1.000	1.010 t (2.844)		-	1.000	1.260 t (86.719)	
	SD	0.000	0.000 p -	0.000	0.111 p 0.005*		-	0.000	0.444 p	0.000*
L_TU	Mean	7.380	1.740 t 4.193	1.240	1.020 t 5.393		-	25.900	8.680 t	7.384
	SD	6.160	1.797 p 0.000*	0.460	0.171 p 0.000*		-	126.949	73.302 p	0.000*
RU1	Mean	0.911	0.865 t 0.762	0.691	0.975 t (9.165)		-	0.026	0.237 t (5.239)	
	SD	0.280	0.277 p 0.446	0.356	0.114 p 0.000*		-	0.087	2.276 p	0.000*
RU2	Mean	0.911	0.887 t 0.422	0.691	1.019 t (10.429)		-	0.041	0.155 t (11.925)	
	SD	0.280	0.259 p 0.673	0.356	0.173 p 0.000*		-	0.158	0.521 p	0.000*
RI1	Mean	1.000	1.000 t -	1.000	0.905 t 9.877		-	1.000	1.002 t (5.091)	
	SD	0.000	0.000 p -	0.000	0.245 p 0.000*		-	0.017	0.042 p	0.000*
RI2	Mean	1.000	1.000 t -	1.000	0.905 t 9.880		-	1.000	1.259 t (86.682)	
	SD	0.000	0.000 p -	0.000	0.244 p 0.000*		-	0.006	0.444 p	0.000*

Table 5. The descriptive result and t-test result in photo, book, bookmark-sharing websites (Random users)

abbr	100 random users (photo)				100 random users (book)				100 random users (bookmark)			
	Yupoo	Yupoo (Rec)	Flickr	Test	Douban	Shelfari	Test	QQ	Del.icio.us	Test		
T	n	0	3466	2779		291		65	16778			
UU	n	7	100	82	1000	1000		78	79			
UT	n	0	80	40	276	39		31	77			
U	n	7	105	86	579	400		78	79			
I	n	613	70844	9296	7938	5452		3368	52877			
NT	Mean	-	32.040	9.970	t	7.629		5.369	8.575	t		
	SD	-	145.913	78.562	p	0.000*		9.325	44.797	p		
NU	Mean	0.000	1057.530	322.290	t	1.597		4.418	1821.165	t		
	SD	-	4177.314	1982.014	p	0.112		15.179	3154.596	p		
NI	Mean	0.000	1.570	2.980	t	(34.232)		0.104	2.721	t		
	SD	-	2.666	3.864	p	0.000*		0.306	7.066	p		
T_U	Mean	0.000	40.830	37.140	t	0.196		0.899	391.582	t		
	SD	-	117.485	142.894	p	0.845		1.809	569.700	p		
I_U1	Mean	87.571	674.850	106.900	t	4.735		45.410	700.975	t		
	SD	-	1066.675	555.637	p	0.000*		50.892	1123.529	p		
I_U2	Mean	-	418.232	141.375	t	1.808		11.226	718.143	t		
	SD	-	908.728	677.871	p	0.074		22.803	1131.985	p		
T_I	Mean	0.000	1.570	2.370	t	(20.804)		0.104	2.719	t		
	SD	-	2.666	3.579	p	0.000*		0.306	7.052	p		
U_I1	Mean	0.011	1.000	1.000	t	(0.981)		1.052	1.047	t		
	SD	-	0.015	0.021	p	0.327		0.333	0.273	p		
U_I2	Mean	-	1.000	1.001	t	(1.579)		1.000	1.047	t		
	SD	-	0.012	0.027	p	0.114		0.000	0.273	p		

Table 5. (continued)

abbr	100 random users (photo)				1000 random users (book)				100 random users (bookmark)			
	Yupoo	Yupoo (Rec)	Flickr	Test	Douban	Shelfari	Test	Test	QQ	Delicio.us	Test	Test
U_T	-	1.240	1.150	t	5.185	1.775	1.230	t	7.509	1.844	t	(15.885)
	-	0.803	0.525	p	0.000*	2.956	0.821	p	0.000*	0.341	p	0.000*
I_T	-	32.040	7.920	t	8.934	3.538	2.510	t	3.350	8.568	t	(0.577)
	-	145.913	56.414	p	0.000*	5.431	4.093	p	0.001*	9.325	p	0.564
T_IU	0.000	1.570	2.980	t	(34.228)	0.715	0.043	t	31.618	2.598	t	(84.481)
	-	2.666	3.863	p	0.000*	2.559	0.367	p	0.000*	0.299	p	0.000*
U_TI	-	1.000	1.260	t	(86.719)	1.006	1.008	t	(0.582)	1.000	t	(0.536)
	-	0.000	0.444	p	0.000*	0.085	0.104	p	0.561	0.000	p	0.592
I_TU	-	25.900	8.680	t	7.384	2.006	2.056	t	(0.187)	4.915	t	0.136
	-	126.949	73.302	p	0.000*	5.050	2.788	p	0.852	4.651	p	0.892
RU1	-	0.026	0.237	t	(5.239)	0.067	0.059	t	1.396	0.090	t	3.468
	-	0.087	2.276	p	0.000*	0.143	0.094	p	0.163	0.180	p	0.001*
RU2	-	0.041	0.155	t	(11.925)	0.067	0.239	t	(10.008)	0.441	t	9.767
	-	0.158	0.521	p	0.000*	0.143	0.324	p	0.000*	0.367	p	0.000*
RI1	-	1.000	1.002	t	(5.091)	0.551	0.550	t	0.054	0.968	t	3.114
	-	0.017	0.042	p	0.000*	0.393	0.420	p	0.957	0.129	p	0.002*
RI2	-	1.000	1.259	t	(86.682)	0.638	0.855	t	(21.136)	1.000	t	120.126
	-	0.006	0.444	p	0.000*	0.376	0.258	p	0.000*	0.000	p	0.000*

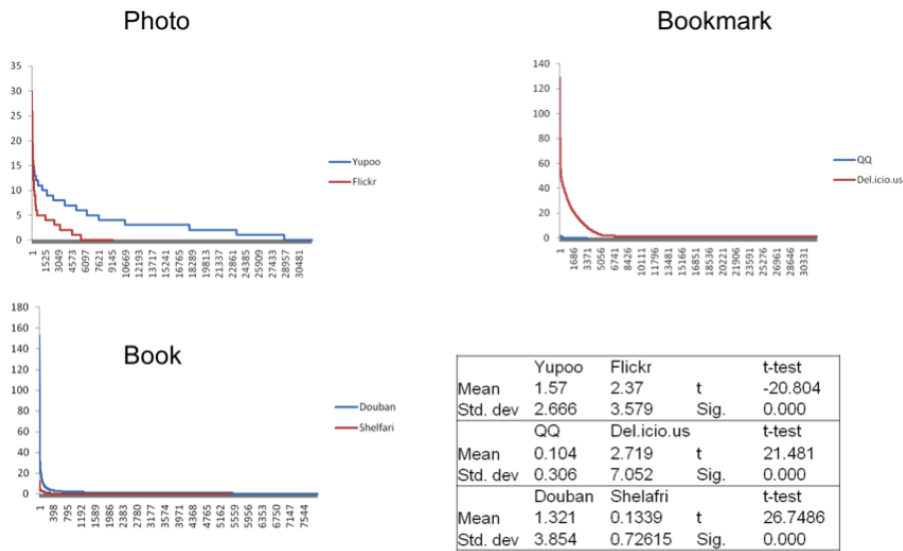


Fig. 1. Less unique tags for each item in Chinese website (Random user)

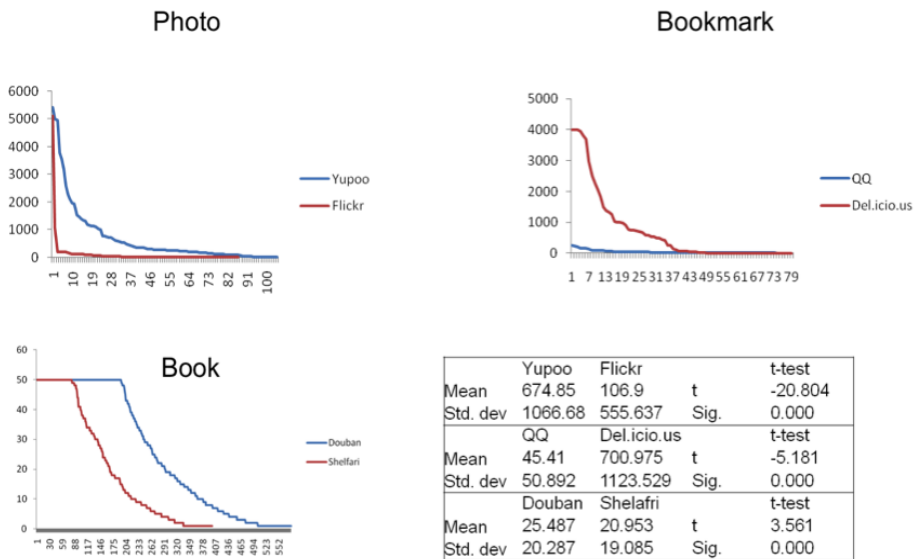


Fig. 2. More number of unique tags for each Chinese user (Random user)

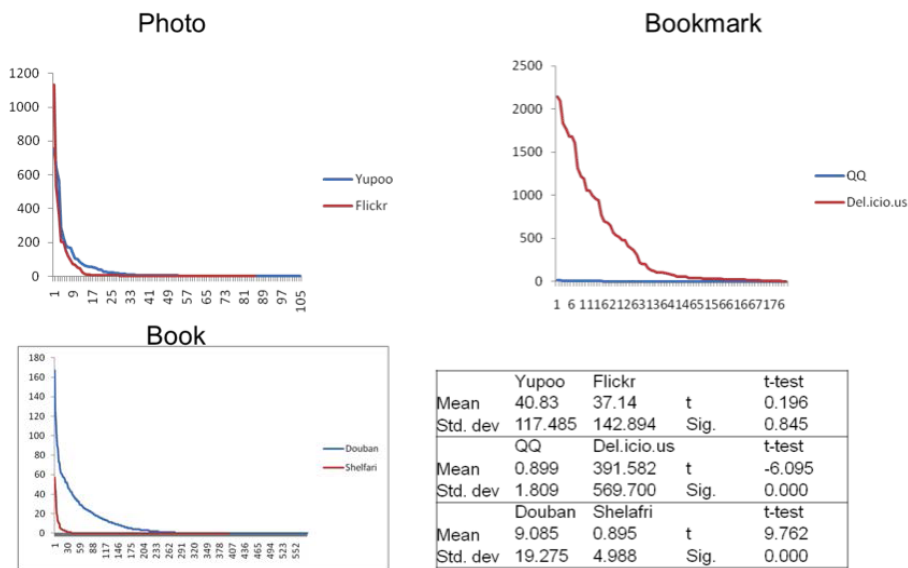


Fig. 3. More number of unique tags for each Chinese user (Random user)

Users use similar tags in photo shared. In recently uploaded photos, Chinese users tend to upload more photos at one time, and they would like to use similar tags for these photos. For example, a user take may photos at Paris and he will upload a lot of photo with tag as Paris on the photo.

Chinese websites users are more willing to share photos, books with tags from his/her history tags and recommendation tag list, but less willing to share bookmarks with tags from his/her history tags and recommendation tag list. This is demonstrated in the value of RU1, RU2, RI1 and RI2.

The result of data analysis demonstrates that most users share less knowledge (photo, book, and bookmark) and fewer users share most knowledge in knowledge sharing websites. Moreover, Chinese culture has effect on the usage pattern in tagging system of knowledge sharing websites; So Chinese users will have a different value in these indices. The reason for this is the restriction of Chinese culture such as modesty. Kurman’s research states that this is the main reason that people’s low of enhancement [11]. Chinese users tend to use tags from her/his history tags, recommendation tags or tags from other users. Chinese will have a higher value in indices RU1, RU2, RI1 and RI2. These indices can reflect the usage pattern of the situation of a knowledge sharing system with a tagging function. This study is helpful for the cultural difference research in knowledge sharing system, and for managing the cross-culture team.

Further research can be developed in three aspects: 1) to analysis the semantic difference of tag of Chinese users; 2) to conduct a questionnaire and interview to study the reason leading to these differences; 3) to collect data from more types of knowledge sharing websites.

References

1. Bhagat, R.S., Kedia, B.L., Harveston, P.D., Triandis, H.C.: Cultural variations in the cross-border transfer of organizational knowledge: An integrative framework. *Academy of Management Review*, 204–221 (2002)
2. Chiu, L.H.: A cross-cultural comparison of cognitive styles in Chinese and American children. *International Journal of Psychology* 7(4), 235–242 (1972)
3. Chow, C.W., Deng, F.J., Ho, J.L.: The openness of knowledge sharing within organizations: a comparative study of the United States and the People's Republic of China. *Journal of Management Accounting Research* 12, 65 (2000)
4. Davenport, T., Prusak, L., Wills, G., Alani, H., Ashri, R., Crowder, R., et al.: *Working knowledge* (1998)
5. Golder, S.A., Huberman, B.A.: Usage patterns of collaborative tagging systems. *Journal of Information Science* 32(2), 198 (2006)
6. Gupta, A.K., Govindarajan, V.: Organizing for knowledge flows within MNCs* 1. *International Business Review* 3(4), 443–457 (1994)
7. Hambrick, D.C., Davison, S.C., Snell, S.A., Snow, C.C.: When groups consist of multiple nationalities: Towards a new understanding of the implications. *Organization Studies* 19(2), 181 (1998)
8. Hofstede, G.: Dimensions of national cultures in fifty countries and three regions. *Expiscations in Cross-Cultural Psychology*, 335–355 (1983)
9. Holden, N.: Knowledge management: raising the spectre of the cross cultural dimension. *Knowledge and Process Management* 8(3), 155–163 (2001)
10. Kipp, M.E.I., Campbell, D.G.: Patterns and inconsistencies in collaborative tagging systems: An examination of tagging practices. *Proceedings of the American Society for Information Science and Technology* 43(1), 1–18 (2006)
11. Kurman, J.: Why is Self-Enhancement Low in Certain Collectivist Cultures? *Journal of Cross-Cultural Psychology* 34(5), 496 (2003)
12. Marlow, C., Naaman, M., Boyd, D., Davis, M.: HT 2006, tagging paper, taxonomy, Flickr, academic article, to read (2006)
13. Marlow, C., Naaman, M., Boyd, D., Davis, M.: HT 2006, tagging paper, taxonomy, Flickr, academic article, to read. Paper Presented at the Proceedings of the Seventeenth Conference on Hypertext and Hypermedia (2006)
14. Nisbett, R.E., Peng, K., Choi, I., Norenzayan, A.: Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review* 108(2), 291 (2001)
15. Shirky, C.: *Ontology is overrated: Categories, links, and tags*. Clay Shirky's Writings About the Internet (2005)
16. Yang, K.: *Chinese personality and its change* (1986)