Proceedings of the Ninth International AAAI Conference on Web and Social Media

Patterns in Interactive Tagging Networks

Yuto Yamaguchi†, **Mitsuo Yoshida**‡, **Christos Faloutsos**§ and **Hiroyuki Kitagawa**†
†University of Tsukuba; ‡Toyohashi University of Technology; §Carnegie Mellon University
yuto_ymgc@kde.cs.tsukuba.ac.jp, yoshida@cs.tut.ac.jp, christos@cs.cmu.edu, kitagawa@cs.tsukuba.ac.jp

Abstract

How do users behave if they can tag each other in social networks? In this paper, we answer this question by studying the *interactive tagging network* constructed by Twitter lists. Twitter lists can be regarded as the tagging process; a user (i.e., tagger) creates a list with a name (i.e., tag) and adds other users (i.e., tagged users) into the list. This tagging network is by nature different from the resource tagging networks (e.g., Flickr and Delicious) because users on this network can tag *each other*. We address the following research questions: (RQ1) What is the common patterns and the difference between the interactive tagging network and the resource tagging networks? (RQ2) Do users tag each other on the interactive tagging network? And if so, to what extent? (RQ3) What is the difference between the two types of relationships on Twitter: *who-tags-whom* and *who-follows-whom*? By quantitatively studying million-scale networks, we found the pervasive patterns across the different tagging networks, and the interactive patterns within the interactive tagging network. This study sheds light on the underlying characteristics of the interactive tagging network, which is relevant to the social scientists and the system designers of the tagging systems.

1 Introduction

How do users behave if they can tag each other in social networks? What is the difference between tagging photos and tagging users? If Smith tags Johnson as *friends*, does Johnson tag Smith back? Same question, if Smith tags Tompson as *sports*, does Tompson tag Smith back? In this paper, we answer these questions by studying the *interactive social tagging* on Twitter [1], where users can tag *each other*.

The social tagging systems like Flickr [2] and Delicious [3] have been studied for years as the novel tools for annotating resources without the central control and the maintained vocabulary (Golder and Huberman 2006; Gupta et al. 2010). Users on the systems are allowed to tag arbitrary resources (e.g., photos) with arbitrary tags as they like. The social tagging systems have attracted a lot of attentions of researchers in this field because of their complex but use-

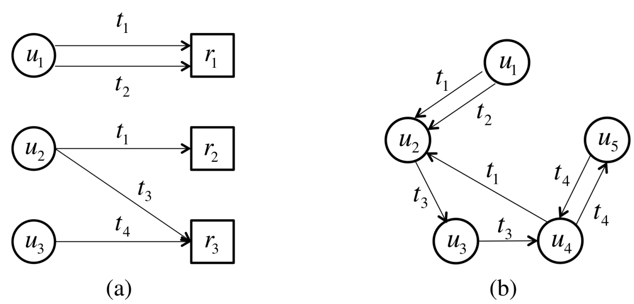

Figure 1: **Two types of tagging networks**: (a) Resource tagging network where users are allowed to tag resources, and (b) Interactive tagging network where users are allowed to tag other users.

ful characteristics. The most important feature of the social tagging systems is that although users behave as they like, the systems produce the well-organized tagging result that is useful as the meta-data of resources (Halpin, Robu, and Shepherd 2007). Researchers have also studied several aspects of the tagging systems: analyzing users' tagging motivations (Heckner, Heilemann, and Wolff 2009; Körner et al. 2010b), identifying the user interests on the tagging systems (Li, Guo, and Zhao 2008; Yin et al. 2011), and investigating the usefulness of the tagging results for the recommender systems (Guan et al. 2010) and the web search (Bao et al. 2007; Heymann, Koutrika, and Garcia-Molina 2008).

Social tagging systems are often modeled as the graphs where users and resources are represented as nodes and these nodes are connected if users tag resources. We call this network the *resource tagging network*. Figure 1(a) shows an example of resource tagging networks. In most cases, users on this network are only allowed to tag resources, in other words, users cannot tag each other.

In contrast to the resource tagging networks, Twitter users can tag other users by creating *Twitter lists*. Creating Twitter lists can be regarded as the tagging process as follows (Wagner et al. 2014): a user (i.e., tagger) creates a list with a name (i.e., tag) and adds other users (i.e., resources) into the list. Such tagging interactions can also be modeled as the graph

[1] https://twitter.com
[2] https://www.flickr.com
[3] https://delicious.com

where users are represented as nodes and they are connected if they tag others. We call this network the *interactive tagging network*. Figure 1(b) shows an example of the interactive tagging networks. The interesting point of this network is that users can be taggers and resources *at the same time*.

The interactive tagging network is worth studying on various aspects. First, it is important for system designers and application developers of tagging systems to identify the differences between the interactive tagging network and the resource tagging networks. For example, the tagging motivation of users is important to design tagging systems because it affects the quality of the tagging results (Körner et al. 2010a). Second, understanding the user behaviors on the interactive tagging network itself is important from the social scientific point of view. To the best of our knowledge, there are very few researches studying the interactive perspective of the Twitter tagging network constructed from Twitter lists (see Section 7 for related work). Third, there are potentially a lot of applications using this *user-tags-user* data, such as user recommendation, ads, and community discovery. Understanding user behaviors on the interactive tagging network helps develop these applications.

Research questions. In this paper, we mainly address the following three research questions:

- **(RQ1) Contrast**: What is the common patterns and the difference between the interactive tagging network and the resource tagging networks?
- **(RQ2) Tagging reciprocity**: Do users tag each other on the interactive tagging network? And if so, to what extent?
- **(RQ3) Tag vs. follow**: What is the difference between the two types of relationships on Twitter: *who-tags-whom* and *who-follows-whom*?

To answer these questions, we study four million-scale networks; three tagging networks on Twitter, Flickr, and Delicious, and a who-follows-whom network on Twitter.

Contributions. Our main contributions and findings are summarized as follows:

- **Pervasive patterns**: We compare the three tagging networks on Twitter, Flickr, and Delicious to answer **RQ1 (Contrast)**. Concretely, our main results include the followings:
  - **Tagging power law**: the interplay among three numbers representing user behaviors follows power law: the number of tags used, the number of resources tagged, and the out-degrees in the tagging network.
  - **Categorizers**: most of Twitter users are *categorizers* (Körner et al. 2010b) who use tags for categorizing resources, which is different from Flickr and Delicious where the population splits into both categorizers and *describers* who use tags for describing resources.
  - **Suspicious user accounts**: there are much more users with suspicious behaviors than those on Flickr and Delicious.

  These findings are derived from Figure 4 where the above three numbers are plotted. We elaborate this figure in Section 4.
- **Interactive patterns**: We study the interactive tagging behaviors of users through investigating the Twitter tagging network and the who-follows-whom network to answer **RQ2 (Tagging reciprocity)** and **RQ3 (Tag vs. follow)**. We mainly find the following patterns/observations in this part:
  - **In-out power law**: the out-degree can be expressed as a power law function of the in-degree.
  - **RM-equation**: the probability of reciprocity can be written as a logarithmic function of the *multiplicity* which is the number of edges between two nodes.
  - **Friendship/subscription tags**: there are (at least) two types of tags: friendship tags that are used for organizing friends, and subscription tags that are used for subscripting posts from famous users.

Outline. The rest of this paper is organized as follows. We first define the terminology and the settings for three RQs in Section 2. Datasets used in this paper are described in Section 3. Three tagging networks on Twitter, Flickr, and Delicious are compared to answer **RQ1 (Contrast)** in Section 4. The interactive tagging perspective of the Twitter tagging network is explored to answer **RQ2 (Tagging reciprocity)** and **RQ3 (Tag vs. follow)** in Section 5. We discuss the implications and limitations of our study in Section 6. We overview the related work from the standpoint of tagging systems, user relationships, and Twitter lists in Section 7. Finally, we conclude the paper in Section 8.

2 Definitions

A *tagging network* G_T can be seen as a set of triplets (u, r, t) which means that user u tags resource r by tag t. Also, a *following network* G_F can be seen as a set of pairs (u, v) which means that user u follows user v. Based on these networks, we answer the research questions in the following settings:

Setting 1 (RQ1 - Contrast).

- **Given**: three tagging networks G_{T1}, G_{T2}, and G_{T3},
- **Find**: the common patterns and the statistical differences behind the three tagging networks.

Setting 2 (RQ2 - Tagging reciprocity).

- **Given**: an interactive tagging network G_T,
- **Find**: the patterns of the tagging reciprocity. By the tagging reciprocity, we mean the form of mutual relationships between users in the tagging network.

Setting 3 (RQ3 - Tag vs. follow).

- **Given**: a tagging network G_T and a following network G_F,
- **Find**: the patterns and the statistical relationships between the tagging and the following.

Formally, a tagging network $G_T = (U, R, T; E)$ is defined as a directed *multigraph* where user nodes $u \in U$ and resource nodes $r \in R$ are allowed to be connected by multiple labeled edges $e \in E \subset U \times R \times T$. Every labeled edge $e = (u, r, t)$ represents that user u tags resource r by tag

Table 1: Details of datasets used for analysis of networks.

	$\lvert U\rvert$	$\lvert R\rvert$	$\lvert E\rvert$	$\lvert T\rvert$
Twitter (tagging)	7,021,966	7,021,966	20,203,951	1,860,400
Twitter (following)	7,021,966	-	432,349,661	-
Flickr 10% (Görlitz, Sizov, and Staab 2008)	32,085	2,840,444	11,534,706	294,209
Delicious 10% (Görlitz, Sizov, and Staab 2008)	52,909	2,620,910	14,091,441	451,925

t. Resource tagging networks are bipertite graphs, while the interactive tagging network is not a bipertite graph because $U = R$. A following network $G_F = (U; E)$ is defined as a directed *simple* graph where user nodes $u \in U$ are connected by just one directed edge $e \in E \subset U \times U$. Every edge represents that user u follows user v on the following network.

3 Data

In this work, we use three tagging networks on Twitter, Flickr, and Delicious, and one following network on Twitter. The details of these networks are shown in Table 1.

Twitter tagging network. We first randomly sampled 1 million *seed users* from the Twitter sample streams[4] on December 2014. And then we collected two kinds of Twitter lists: lists that are created by the seed users, and lists that the seed users are added into as list members. Based on the collected Twitter lists, we construct the Twitter tagging network as follows: if user u makes a list with a name t and adds user v into the list, we make an edge from u to v with tag t. Consequently, the Twitter tagging network has 7,021,966 nodes, 20,203,951 edges, and 1,860,400 tags.

Twitter following network. The following network on Twitter, which is constructed from the who-follows-whom relationships, can be directly collected using Twitter API. We collected following and followed edges starting from the same 1 million seed users on the Twitter tagging network. To study exactly the same users as the tagging network, we use the only edges among 7,021,966 users on the tagging network. As a result, the Twitter following network has 7,021,966 nodes and 432,349,661 edges. Note that the following network is much denser than the tagging network.

Flickr. The Flickr tagging network is obtained from (Görlitz, Sizov, and Staab 2008), which contains 319,686 users, 28,153,045 resources, 1,607,879 tags, and 112,900,000 edges as a whole. We randomly sampled 10% of all users to adjust the scale of the network to the Twitter tagging network. Consequently, the Flickr tagging network has 32,085 users, 2,840,444 resources, 11,534,706 edges, and 294,209 tags.

Delicious. The Delicious tagging network is also obtained from (Görlitz, Sizov, and Staab 2008), which contains 532,924 users, 17,262,671 resources, 2,481,103 tags, and 140,126,555 edges as a whole. We randomly sampled 10% of all users in the same reason as Flickr. As a consequence, the Delicious tagging network has 52,909 users, 2,620,910 resources, 14,091,441 edges, and 451,925 tags.

[4]https://dev.twitter.com/streaming/reference/get/statuses/sample

Reproducibility. Our datasets of Twitter tagging and following networks are available at http://dx.doi.org/10.5281/zenodo.16267, and our code is also available at http://github.com/yamaguchiyuto/icwsm15. We can download Flickr and Delicious at the authors' web page of (Görlitz, Sizov, and Staab 2008).

4 Contrast - (RQ1)

In this section, we compare the Twitter tagging network with the Flickr and the Delicious tagging networks. We are interested in the pervasive patterns behind the user behaviors on all three tagging networks. To study the *one-way* resource tagging perspective of the Twitter tagging network, we use only the out-going edges from the 1 million seed users, which amounts to 11,115,405 edges (55.0%).

4.1 Global statistics

Firstly, we study the global statistics of tagging networks to compare the characteristics of the tagging results, namely, in-/out-degree distributions and multiplicity distributions. Multiplicity is the number of edges (i.e., tags) from a user to a resource, which means how many times users tag the same resources.

In-/out-degree distributions Is there any statistical difference among the degree distributions of three tagging networks? Figure 2 shows the log-log plots of in-/out-degree distributions of three tagging networks. The x-axes of all plots indicate the in-/out-degrees. The y-axes of the upper three figures indicate the number of nodes with the corresponding x-values, while those of the lower three figures indicate the CCDF values.

It is observed that all distributions are long-tailed distributions, especially, in-degree distributions seem to follow *power-law distributions* (Faloutsos, Faloutsos, and Faloutsos 1999). We also observe in the CCDF plots that the slope of the in-degree distributions on Flickr is steeper than those on Twitter and Delicious as shown by the black line, which indicates that Twitter and Delicious are *broad folksonomy* (Helic et al. 2012) while Flickr is a *narrow folksonomy*. A *broad folksonomy* is a tagging network where a lot of users tag the same resources, while a *narrow folksonomy* is a tagging network where one or a few users tag one resource, especially their *own resources*.

Multiplicity How many times do users tag the same resource? Figure 3 shows the log-log plots of multiplicity distributions, where the x-axes indicate the multiplicity, while

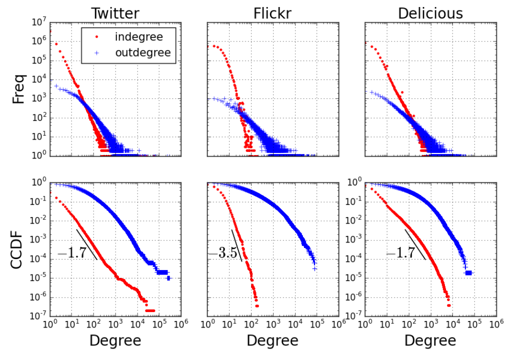

Figure 2: **Contrast - degrees**: Log-log plots of the in-/out-degree distributions. Lower figures show the CCDF plots. Twitter and Delicious are broad folksonomy where a lot of users tag the same resources, while Flickr is a narrow folksonomy where one or a few users tag one resource, which is shown by the slopes of the CCDF of the in-degree distributions.

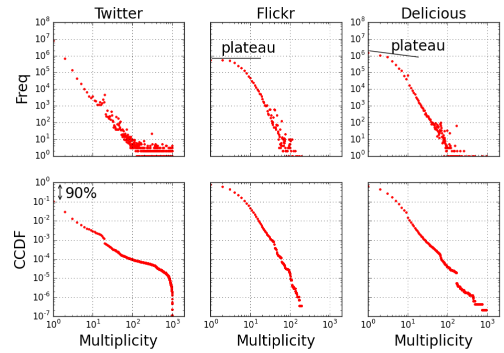

Figure 3: **Contrast - multiplicity**: Log-log plots of multiplicity distributions. Lower three figures show the CCDF plots. A lot of user-resource pairs on Twitter (90%) have only one edge between them ($x = 1$), while those on Flickr and Delicious are less (10%), which is also shown by the plateaus on Flickr and Delicious. This result indicates that Twitter users tag one resource just once in most cases, while Flickr and Delicious users are more likely to tag the same resources multiple times.

the y-axes indicate the frequency and the CCDF. Each point corresponds to a pair of a user and a resource. It is shown that all three distributions are long-tail distributions, meaning there are a small number of pairs with extremely large multiplicities. Note that there is a cut-off at $x = 1,000$ on Twitter because users are not allowed to create over 1,000 lists. In addition, we can see a spike at $x = 20$ on Twitter, which is because of the past limitation that users cannot create over 20 lists.

From the CCDF plots, we observe that the ratio of the user-resource pairs with $x = 1$ on Twitter (90%) is much larger than that of Flickr and Delicious (10%), which is also shown by the plateaus on Flickr and Delicious in the upper figures. This can be explained by the tagging motivations of users. It is said that there are mainly two types of users with different tagging motivations: *categorizers* and *describers* (Körner et al. 2010b; 2010a). Categorizers tend to tag one resource *just once* (i.e., $x = 1$) because they want to explicitly categorize resources with a small number of tags for later browsing, while describers tend to tag one resource *multiple times* (i.e., $x > 1$) because they want to describe resources in detail for later retrieval. Most of Twitter users tag the same resource just once, indicating majorities of Twitter users are categorizers.

There are a large number of Twitter users with large multiplicities (e.g., over 100). These users are likely to be spam user accounts because regular users are not motivated to tag the same resource so many times. This issue about suspicious behaviors is discussed in the next subsection more carefully.

4.2 Individual behaviors

Is there any pervasive pattern behind the individual behaviors on all three tagging networks? Next we investigate the tagging behaviors of individual users in detail. Let $T_u \subset T$ be the set of tags user u used, $R_u \subset R$ be the set of resources user u tagged, and d_u be the out-degree of user u. The interplay of these three measures tells us how users behave on the tagging networks. For example, categorizers tend to tag a lot of resources with a limited number of tags (i.e., $|R_u| >> |T_u|$).

Figure 4 show the scatter plots of all combinations of two out of three measures. Each point corresponds to one user. Black stars show the mean values corresponding to the log-bins, and the red line shows the fitted curve of the mean values by the least square error.

Strikingly, these data are all explained very well by the *power law function* except for the deviation of the mean values observed on Twitter[5]. This means that there is the pervasive pattern behind the user behaviors across the networks. The exponent and the intercept tell us much about the user behaviors on these tagging networks. For example, on the middle three figures, although all the exponents are close to one, the intercepts are different from each other, indicating that the average number of tags assigned to one resource is different (1.11 on Twitter, 2.33 on Flickr, and 2.96 on Delicious).

These all plots indicate there are a lot of categorizers on Twitter, while the user population splits into the categorizers and describers on Flickr and Delicious. It is shown from these figures that there are many points with $|R_u| >> |T_u|$, $|R_u| = d_u$, and $d_u >> |T_u|$ on Twitter. If user u tags a lot of resources with a small number of tags, the values $|R_u|$ and d_u are much larger than $|T_u|$, which indicates that user u is a categorizer who wants to maintain its vocabulary size small. Moreover, user u with $|R_u| = d_u$ does not tag the same resource more than once, which also indicates that the

[5] We ignore the deviations observed on Twitter when fitting to better capture the trends

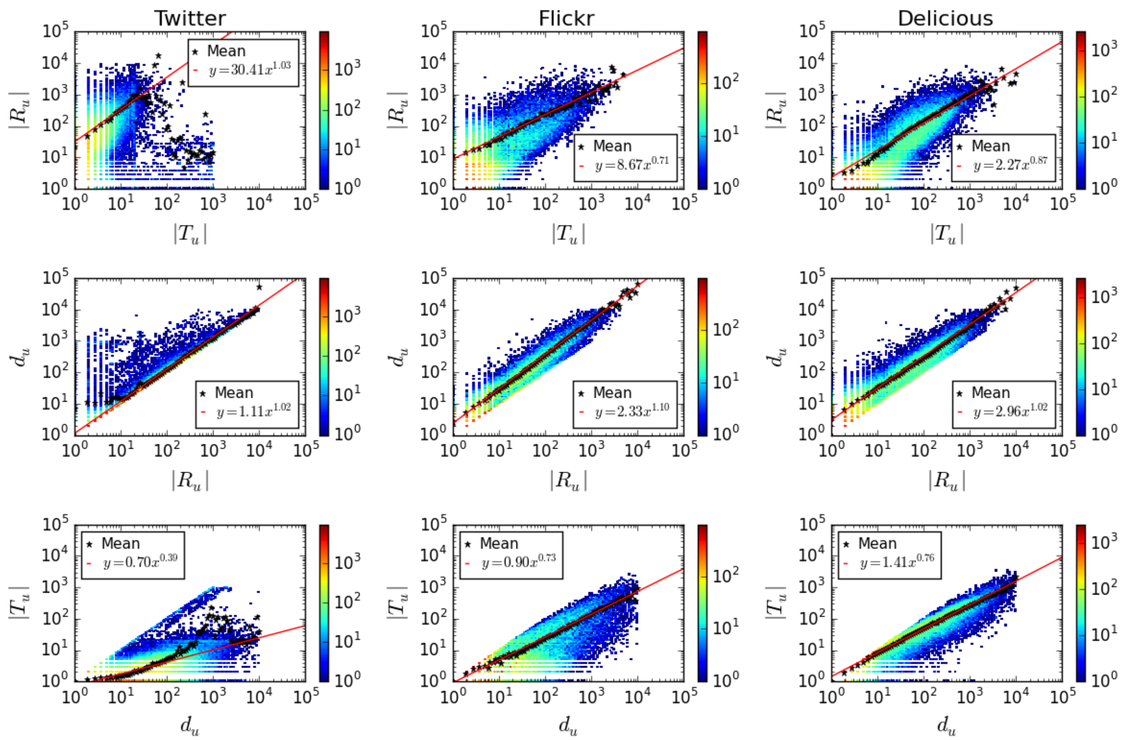

Figure 4: **Contrast - behaviors**: Scatter plots of the individual behaviors on three tagging networks. Each point represents one user. Black stars show the mean values corresponding to the log-bins, and red lines show the fitted power law curve by the least square error. There are much more categorizers who have values $|R_u| = d_u >> |T_u|$ on Twitter than those on Flickr and Delicious. Also, we can observe unnatural point clouds on Twitter that have values $|T_u| = d_u >> |R_u|$.

user wants to explicitly categorize them. This is understandable result because Twitter lists are originally introduced for organizing posts from other users (Kim et al. 2010).

As we can clearly see, there are unnatural point clouds in the scatter plots of Twitter. Specifically, we observe three point clouds on Twitter: (a) $10^2 \le |T_u| \le 10^3$ and $10^0 \le |R_u| \le 10^2$ in the upper-left figure, (b) $10^0 \le |R_u| \le 10^1$ and $10^2 \le d_u \le 10^3$ in the middle-left figure, and (c) $10^2 \le d_u \le 10^3$ and $10^2 \le |T_u| \le 10^3$ in the lower-left figure. These points illustrate the existence of the suspicious user accounts that use an extremely large number of tags, and tag a few resources many times ($|T_u| = d_u >> |R_u|$). We will discuss this issue in Section 6.

4.3 Findings

The patterns/observations we found in this section are summarized as follows:

Pattern 1 (Tagging power law). *On all three tagging networks, the interplay among the three measures of user behaviors follows the power law:*

$$|R_u| = A_1 \cdot |T_u|^{\alpha_1} \quad (1)$$
$$d_u = A_2 \cdot |R_u|^{\alpha_2} \quad (2)$$
$$|T_u| = A_3 \cdot d_u^{\alpha_3} \quad (3)$$

All the constants on three tagging networks are illustrated in Figure 4.

Observation 1 (Categorizers). *Most of Twitter users are* categorizers *who use tags for categorizing resources, which is different from Flickr and Delicious where the population splits into categorizers and* describers *who use tags for describing resources (Figures 3 and 4).*

Observation 2 (Suspicious users). *There are a lot of suspicious user accounts on Twitter (Figure 4).*

Observation 3 (Broad folksonomy). *The Twitter tagging network is a* broad folksonomy *where a lot of users tag the same resource, which is similar to Delicious but different from Flickr (Figure 2).*

5 Interactive patterns

In this section, we study the interactive tagging perspectives of the Twitter tagging network, where users can tag each other. To answer **RQ2 (Tagging reciprocity)**, we first explore the tagging reciprocity of users in the Twitter tagging network. Next we analyze the relationship between the tagging and following behaviors on Twitter to answer **RQ3 (Tag vs. follow)**. Note that in this section we use the whole Twitter tagging network described in Table 1.

5.1 Tagging reciprocity - (RQ2)

The tagging reciprocity plays a crucial role to study the interactive tagging network because users can tag other users and can be tagged by them at the same time.

In-/out-degree correlation If a user is actively tagged by other users, is that user also active for tagging? In other words, is there any correlation between in-degrees and out-degrees on the tagging network? Figure 5 shows the log-binned plot of the result. The x-axis shows in-degree values, while the y-axis indicates the mean out-degree values of users of the corresponding bins. Blue dots represent the data, and the red line shows the curve fitted by the least square error.

The result demonstrates that users who are tagged many times tend to tag a lot of other users. One possible explanation of this is that users tagged a lot may perceive the usage of the Twitter lists and start to use lists. Another explanation is that users use tags (i.e., Twitter lists) as the communication tool. If it is the case, users who are tagged by their friends are likely to tag their friends back, which is indeed observed in the next two subsections.

The interesting point is that the fitted curve shown as the red line is a *power law* function with the exponent < 1. Another interesting point is that although the exponent is less than 1, it is still growing after $x = 100$. This is surprising because it indicates that famous users who are tagged more than 100 times (only 0.05% of all population) are more likely to be active for tagging than regular users. By looking into the data, we observe that, for example, famous news accounts that are tagged a lot tend to make lists and add other informative accounts into their lists. Although this is the first study to show the in-/out-pattern described as a power law function, this *correlation* itself between in-degrees and out-degrees is also observed in the Twitter following network (Java et al. 2007) and in the Facebook wall (Saez-Trumper, Nettleton, and Baeza-Yates 2011), which suggests that the correlation between in-coming and out-going actions is *pervasive* in social networks.

Reciprocity vs. multiplicity If user u tags v, how likely does v tag u back? Does the reciprocity probability increase if users are tagged multiple times by the same user? To study the interplay between the reciprocity and the multiplicity, we compute the **RP@m**, which is the reciprocity probability as a function of the multiplicity values.

Figure 6 shows the log-binned plot of the result, where the x-axis indicates the multiplicity in the log scale, while the y-axis shows the RP@m values of the corresponding bins. Blue dots represent the data, and the red line shows the fitted curve by the least square error. There are two trends in this figure; RP@m clearly increases logarithmically before $x = 20$ as shown by the fitted curve, but it goes down after that and reaches $y = 0$. The result indicates that users are more likely to tag back other users if they are tagged more, but less likely to tag back if they are tagged *too much*. This implies that friends tag each other multiple times (but up to 20), while spam accounts unidirectionally tag the targets too much but they never get reciprocated.

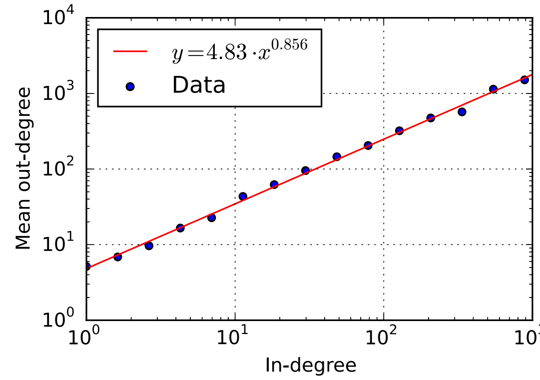

Figure 5: **In-out power law**: Log-binned plot of the in-/out-degrees. The y values indicate the mean out-degrees of the corresponding bins. The red curve represents the fitted curve by the least square error. Users who are actively tagged by other users tend to be active for tagging. This trend is also true even for famous users who are tagged more than 100 times (0.05% of all population).

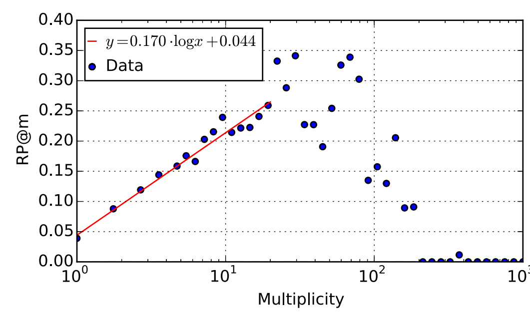

Figure 6: **RP@m**: Log-binned plot of the reciprocity probability as a function of the multiplicity (RP@m). The red curve represents the fitted curve by the least square error. There are two trends; RP@m clearly increases logarithmically before $x = 20$ as shown by the fitted curve, but it goes down after that and reaches $y = 0$, which indicates that users are more likely to tag other users back if they are tagged more, but less likely to reciprocate if they are tagged too much.

Reciprocity vs. tags Which kind of tags are more likely to be reciprocated? In this subsection we compute the **RP@t**, which is the reciprocity probability as a function of tags. That is, RP@t indicates how likely a user tags back another if the latter tags the former by tag t.

Table 2 shows the RP@t values of the top 20 frequent tags and their *lift* which is the ratio of RP@t against the overall reciprocity probability ($= 0.046$). Tags in this table are sorted by their occurrence frequency. Interestingly, tags related to the friendship (e.g., *friends* and *faves*) have large lift values. In contrast, tags related to the information topics (e.g., *news*, *music*, and *sports*) have small lift values. For ex-

Table 2: **RP@t** of the top 20 frequent tags. RP@t is the reciprocity probability as a function of tags. Lift is the ratio of RP@t to the overall reciprocity probability ($= 0.046$). The representative friendship tag *friends* has 15 times larger value of RP@t than that of the representative subscription tag *news*.

Tags	Freq.	RP@t	Lift
news	64,261	0.005	0.105
my-favstar-fm-list	49,743	0.132	2.897
music	23,473	0.011	0.247
bot	22,406	0.004	0.079
friends	21,723	0.070	1.537
sports	18,962	0.006	0.130
noticias	18,500	0.003	0.067
media	14,120	0.013	0.283
amigos	13,575	0.036	0.785
politics	9,345	0.013	0.291
list	8,301	0.054	1.193
tech	8,249	0.013	0.295
social-media	7,701	0.034	0.752
football	6,788	0.020	0.435
entertainment	5,742	0.007	0.143
travel	5,669	0.025	0.548
sport	5,651	0.005	0.100
deportes	5,644	0.002	0.038
business	5,389	0.017	0.377
faves	5,187	0.131	2.880

Table 3: **Tag-follow disagreement**: Approximately 40% of user relationships on the tagging network is missing on the following network.

Type	Prob.
w/ direction	0.590
w/o direction	0.621

ample, friendship tag *friends* has 15 times larger RP@t value than subscription tag *news*.

This result implies that there are (at least) two types of tags: *friendship* tags and *subscription* tags. The friendship tags seem to be used for maintaining friendships, while the subscription tags seem to be used for subscribing posts from famous users related to the corresponding topics. In the next section, we look at this point from another angle, namely, the relationship between the tagging and following.

5.2 Tag vs. follow - (RQ3)

In this section, we answer the **RQ3 (Tag vs. follow)** through the analyses of the relationship between the tagging and the following. To the best of our knowledge, this is the first study to analyze the overlaid network of the tagging and following on Twitter.

Tag-follow agreement Do the tagging network and the following network agree with each other with regard to their structures? In other words, how likely does user u follow v on the following network if u tags v on the tagging network?

Table 3 shows the result, where two types of probabilities are reported. The upper is the probability that user u follows v conditioned by the fact that u tags v on the tagging network, that is, it considers the edge direction. The lower is the probability that u follows v or v follows u conditioned by the fact that u tags v, that is, it does not consider the direction of following edges. It is shown that approximately 40% of user relationships on the tagging network is missing on the following network, in other words, the tagging and the following networks *disagree* with each other in their structures. This is surprising because the number of edges in the tagging network is much less than the number of edges in the following network.

Table 4: **Forward and mutual ratio** of the top 20 frequent tags. The overall ratios are Mutual $= 0.549$ and Forward $= 0.213$.

Tags	Mutual	Forward
news	0.096	**0.479**
my-favstar-fm-list	**0.561**	0.132
music	0.213	**0.386**
bot	**0.351**	0.132
friends	**0.578**	0.210
noticias	0.092	**0.555**
sports	0.086	**0.536**
media	0.196	**0.443**
amigos	0.327	**0.420**
politics	0.265	**0.345**
list	**0.415**	0.232
tech	0.166	**0.435**
social-media	**0.415**	0.191
football	0.140	**0.445**
entertainment	0.118	**0.418**
travel	**0.309**	0.286
sport	0.085	**0.513**
deportes	0.087	**0.491**
business	**0.324**	0.257
faves	**0.608**	0.128

Following patterns If user u tags v by tag t, does u unidirectionally follows v? Or do u and v mutually follow each other? The former pattern is referred to as *forward*, while the latter is referred to as *mutual*. Here we compute the forward ratio and the mutual ratio. The forward ratio is the probability that user u unidirectionally follows user v under the condition that u tags v, while the mutual ratio is the probability that u and v follow each other under the condition that u tags v.

Table 4 shows the mutual and the forward ratios of the top 20 frequent tags. We can see that friendship tags (e.g., *friends* and *faves*) have large mutual ratio, while subscription tags (e.g., *news* and *sports*) have small mutual ratio. This result means that users are likely to be mutual friends on the following network if one user tags another by friendship tags. In contrast to friendship tags, even if a user tags another by subscription tags, the latter will not follow the former.

Figure 7 plots these ratios of the top 30 frequent tags. The x-axis indicates the mutual ratio, while the y-axis shows the forward ratio. Each blue dot represents each tag, and the red line shows the fitted line by the least square error. We can

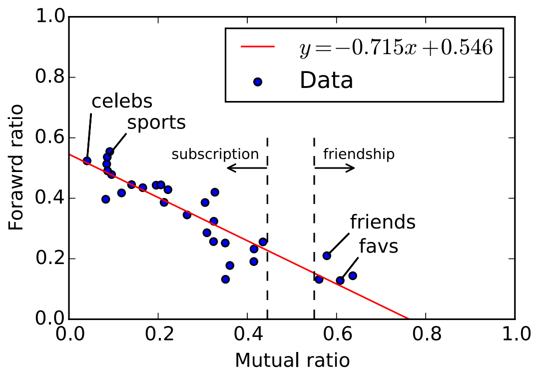

Figure 7: **Gap in behaviors**: Plot of the forward and the mutual ratios of the top 30 frequent tags. Each blue dot represents each tag, and the red line shows the fitted line by the least square error. There is a clear gap around $x = 0.5$, which separates friendship tags and subscription tags.

see that the data is well explained by the fitted line. Interestingly, we observe that there is a clear gap around $x = 0.5$, which indicates that friendship tags and subscription tags are clearly separated in terms of the following patterns. Indeed, representative subscription tags *celebs* and *sports* are on the left-hand side, while friendship tags *friends* and *favs* are on the right-hand side.

5.3 Findings

In summary, we found the following pattens/observations through investigating the tagging and following networks on Twitter:

Pattern 2 (In-out power law). *The tagging activeness (i.e., out-degree) can be written as a* power-law *function of the number of times users are tagged (i.e., in-degree) as follows:*

$$d_{out} = A \cdot d_{in}^{\alpha} \tag{4}$$

where d_{out} and d_{in} are the out-degree and the in-degree, respectively. Specifically, the constants on the Twitter tagging network are $A = 4.83$ and $\alpha = 0.856$. Interestingly, this trend is also applicable for famous users (0.05% of all population) who are tagged more than 100 times (Figure 5).

Pattern 3 (RM-equation). *User u is more likely to tag user v back if v tags u more, which is expressed as the following equation:*

$$r = \beta \cdot \log m + B \tag{5}$$

where r is the reciprocity probability, and m is the multiplicity. This trend is in the region $m < 20$. It is reversed after that and reaches $r = 0$, which may be caused by anomalous users. The constants are $\beta = 0.170$ and $B = 0.044$ on Twitter (Figure 6).

Pattern 4 (FM-equation). *There is a linear relationship between the forward ratio and the mutual ratio as follows:*

$$m = \gamma \cdot f + C \tag{6}$$

where m is the mutual ratio and f is the forward ratio. The forward ratio denotes how likely user u unidirectionally follows user v if u tags v, while the mutual ratio denotes how likely u and v follow each other if u tags v. On Twitter, the constants are $\gamma = -0.715$ and $C = 0.545$ (Figure 7).

Observation 4 (Friendship/subscription tags). *There are two types of tags:* friendship *and* subscription *tags. Friendship tags are used for organizing friends, while subscription tags are used for subscripting posts from famous users. There is a clear gap around $x = 0.5$ in Figure 7, which separates these two types of tags.*

Observation 5 (Tag-follow disagreement). *A large part (40%) of user relationships on the tagging network are not present on the following network in spite of the fact that the number of edges on the following network is much larger than that of the tagging network (Table 3).*

6 Discussion

We have revealed several underlying patterns of the interactive tagging network so far. In this section, we discuss the additional implications and the limitations of this work.

Additional implications. As we mentioned in the earlier section, we observe that there are a large number of suspicious user behaviors on Twitter. For example, we can see the extremely large multiplicity values (Figure 3) and the unnatural point clouds (Figure 4). These results imply that there are a lot of *spam accounts* on Twitter compared to Flickr and Delicious. We indeed observe many spams that create a lot of lists with meaningless names (e.g., 500 lists named from *list-1* to *list-500*), which may be made by scripts. Although understanding the motivations of these spam accounts is an open problem, we can manually spot them by looking into the abnormal point clouds in Figure 4, which results in more quality folksonomy outcome.

From the result of the tag-follow disagreement, we would say that it is not enough to analyze only the following network for studying user interactions on Twitter. Approximately 40% of Twitter users do not follow other users if they create a list and add other users into the list, especially if they use *subscription tags* (Table 4). This means that we are *ignoring* a large part of subscription or information collection relationships between famous users and ordinary users on Twitter if we only study the following network.

Limitations. First, we studied the behaviors of *only* users who use Twitter lists. Hence, the following question remains as an open question: *what is the difference between behaviors of users who use Twitter lists and those who do not?* Second, we simply modeled the Twitter lists as the tagging network where an edge from user u to user v by tag t means that u makes a list with name t and adds v as a member of the list. However, there is another rich feature of the Twitter lists. Twitter users can *subscribe* lists that are created by other users. By analyzing this rich feature of the Twitter lists, we may be able to evaluate the quality and the credibility of the list creators. Third, we focused on the *network analysis* of the tagging networks in this paper, meaning that the semantics of tags itself is not investigated. Although these limitations may lead to several interesting open questions about the Twitter list itself, we believe our study discovered important and general inter-network and within-network patterns of the interactive tagging network.

7 Related Work

We survey the related work in the areas of social tagging, user relationships, and Twitter lists.

Social tagging. There are numerous social tagging systems like Flickr and Delicious, which have been studied for years. (Golder and Huberman 2006) studied several aspects of Delicious, which is the representative social tagging system. For example, they reported that the growth of the number of tags users used reflects the transition of the users' interests. (Halpin, Robu, and Shepherd 2007) showed that the distribution of use of tags for popular web pages on Delicious follow the power law distribution, which often emerge in the complex systems. They also proposed a generative model to understand the dynamics of tag usage. (Li, Guo, and Zhao 2008) and (Yin et al. 2011) studied the user interests on the social tagging systems based on the assumption that users tend to tag the resources they are interested in. They both proposed different methods to model the user interests. The motivation of users for social tagging has been investigated in a line of research (Heckner, Heilemann, and Wolff 2009; Strohmaier, Körner, and Kern 2010; Körner et al. 2010a; 2010b). It has been shown that the tagging motivation differs not only across the social tagging systems, but also within the tagging systems. Although these studies have successfully revealed several characteristics underlying the resource tagging networks, they are different from our work in that we focus on the interactive tagging network where users can tag each other.

User relationships. We can associate with other people in various kinds of social networks. (Adamic et al. 2011) performed the qualitative and quantitative survey of the relationships between two types of user interactions: friendship and trust. They concluded that close friendship includes high-level trust, while the latter can be achieved without the former. (Leskovec, Huttenlocher, and Kleinberg 2010b) studied the interplay between the positive and the negative links in terms of the *social balance* and the *social status*. They also proposed a method for predicting the sign of links based on their findings in (Leskovec, Huttenlocher, and Kleinberg 2010a). (Hopcroft, Lou, and Tang 2011) investigated the reciprocity on the Twitter following network and proposed a method for predicting the reciprocity. (Barbieri, Bonchi, and Manco 2014) proposed a generative model to explain why users follow other users using hashtags on Twitter. These research differs from ours in that they do not focus on the explicit tags exchanged between users.

Twitter lists. (Kim et al. 2010) studied the basic characteristics of Twitter lists, and reported that Twitter users can be classified into topics using Twitter lists. Twitter lists have been used to achieve several applications, such as inferring user demographics (Yamaguchi, Amagasa, and Kitagawa 2011), analyzing user communications in different categories (Wu et al. 2011), and recommendations (Rakesh et al. 2014). To the best of our knowledge, (Zhao and Ram 2011) is the only study that analyzes the user interactions in the Twitter tagging network. However, it differs from ours in that it focuses on just the triadic closure.

8 Conclusion

In this work, we studied the interactive tagging network on Twitter, which is constructed by Twitter lists. Users on the interactive tagging network can tag each other, in contrast to the resource tagging networks (e.g., Flickr and Delicious). Our contributions in this paper are summarized as follows:

- **Pervasive patterns**: We compare three tagging networks on Twitter, Flickr, and Delicious to find the pervasive patterns and the statistical differences among them (Section 4). Our main results include: *tagging power law* (Pattern 1; Equations 1, 2, and 3), *categorizers* (Observation 1), *suspicious users* (Observation 2), and *broad folksonomy* (Observation 3).
- **Interactive patterns**: We investigated the interactive tagging perspective of the Twitter tagging network (Section 5). We found several patterns/observations of the user interactions, namely, *in-out power law* (Pattern 2; Equation 4), *RM-equation* (Pattern 3; Equation 5), *FM-equation* (Pattern 4; Equation 6), *friendship/subscription tags* (Observation 4), and *tag-follow disagreement* (Observation 5).

Our future work includes the analysis of the information diffusion pattern using the tagging network where edges are labeled by tags. If we know the reason why user A follows B based on the tags A assigned to B (e.g., A is interested in B's tweets about politics), we can identify the *topic-specific* information flow on the network. In addition, we plan to develop a generative model of the interactive tagging between users, which leads to a deeper understanding of the user behaviors.

Acknowledgements

: This research was partly supported by the program *Research and Development on Real World Big Data Integration and Analysis* of the Ministry of Education, Culture, Sports, Science and Technology, Japan, and by JSPS KAKENHI, Grant-in-Aid for JSPS Fellows #242322. This material is also based upon work supported by the National Science Foundation (Grants No. IIS-1217559 and CNS-1314632), and by the Army Research Laboratory (under Cooperative Agreement Number W911NF-09-2-0053). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation, or other funding parties. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

References

Adamic, L. A.; Lauterbach, D.; Teng, C.; and Ackerman, M. S. 2011. Rating friends without making enemies. In *ICWSM*.

Bao, S.; Xue, G.; Wu, X.; Yu, Y.; Fei, B.; and Su, Z. 2007. Optimizing web search using social annotations. In *Proceedings of the 16th international conference on World Wide Web*, 501–510. ACM.

Barbieri, N.; Bonchi, F.; and Manco, G. 2014. Who to follow and why: link prediction with explanations. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 1266–1275. ACM.

Faloutsos, M.; Faloutsos, P.; and Faloutsos, C. 1999. On power-law relationships of the internet topology. In *SIGCOMM*, 251–262.

Golder, S. A., and Huberman, B. A. 2006. Usage patterns of collaborative tagging systems. *Journal of information science* 32(2):198–208.

Görlitz, O.; Sizov, S.; and Staab, S. 2008. Pints: peer-to-peer infrastructure for tagging systems. In *IPTPS*, 19.

Guan, Z.; Wang, C.; Bu, J.; Chen, C.; Yang, K.; Cai, D.; and He, X. 2010. Document recommendation in social tagging services. In *Proceedings of the 19th international conference on World wide web*, 391–400. ACM.

Gupta, M.; Li, R.; Yin, Z.; and Han, J. 2010. Survey on social tagging techniques. *ACM SIGKDD Explorations Newsletter* 12(1):58–72.

Halpin, H.; Robu, V.; and Shepherd, H. 2007. The complex dynamics of collaborative tagging. In *Proceedings of the 16th international conference on World Wide Web*, 211–220. ACM.

Heckner, M.; Heilemann, M.; and Wolff, C. 2009. Personal information management vs. resource sharing: Towards a model of information behavior in social tagging systems. In *ICWSM*.

Helic, D.; Körner, C.; Granitzer, M.; Strohmaier, M.; and Trattner, C. 2012. Navigational efficiency of broad vs. narrow folksonomies. In *Proceedings of the 23rd ACM conference on Hypertext and social media*, 63–72. ACM.

Heymann, P.; Koutrika, G.; and 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. ACM.

Hopcroft, J.; Lou, T.; and Tang, J. 2011. Who will follow you back?: reciprocal relationship prediction. In *Proceedings of the 20th ACM international conference on Information and knowledge management*, 1137–1146. ACM.

Java, A.; Song, X.; Finin, T.; and Tseng, B. L. 2007. Why we twitter: An analysis of a microblogging community. In *WebKDD/SNA-KDD*, 118–138.

Kim, D.; Jo, Y.; Moon, I.-C.; and Oh, A. 2010. Analysis of twitter lists as a potential source for discovering latent characteristics of users. In *ACM CHI Workshop on Microblogging*.

Körner, C.; Benz, D.; Hotho, A.; Strohmaier, M.; and Stumme, G. 2010a. Stop thinking, start tagging: tag semantics emerge from collaborative verbosity. In *Proceedings of the 19th international conference on World wide web*, 521–530. ACM.

Körner, C.; Kern, R.; Grahsl, H.-P.; and Strohmaier, M. 2010b. Of categorizers and describers: An evaluation of quantitative measures for tagging motivation. In *Proceedings of the 21st ACM conference on Hypertext and hypermedia*, 157–166. ACM.

Leskovec, J.; Huttenlocher, D.; and Kleinberg, J. 2010a. Predicting positive and negative links in online social networks. In *Proceedings of the 19th international conference on World wide web*, 641–650. ACM.

Leskovec, J.; Huttenlocher, D.; and Kleinberg, J. 2010b. Signed networks in social media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1361–1370. ACM.

Li, X.; Guo, L.; and Zhao, Y. E. 2008. Tag-based social interest discovery. In *Proceedings of the 17th international conference on World Wide Web*, 675–684. ACM.

Rakesh, V.; Singh, D.; Vinzamuri, B.; and Reddy, C. K. 2014. Personalized recommendation of twitter lists using content and network information. In *ICWSM*.

Saez-Trumper, D.; Nettleton, D. F.; and Baeza-Yates, R. A. 2011. High correlation between incoming and outgoing activity: A distinctive property of online social networks? In *ICWSM*.

Strohmaier, M.; Körner, C.; and Kern, R. 2010. Why do users tag? detecting users' motivation for tagging in social tagging systems. In *ICWSM*.

Wagner, C.; Singer, P.; Strohmaier, M.; and Huberman, B. A. 2014. Semantic stability in social tagging streams. In *Proceedings of the 23rd international conference on World wide web*, 735–746. International World Wide Web Conferences Steering Committee.

Wu, S.; Hofman, J. M.; Mason, W. A.; and Watts, D. J. 2011. Who says what to whom on twitter. In *Proceedings of the 20th international conference on World wide web*, 705–714. ACM.

Yamaguchi, Y.; Amagasa, T.; and Kitagawa, H. 2011. Tag-based user topic discovery using twitter lists. In *ASONAM*, 13–20.

Yin, D.; Hong, L.; Xue, Z.; and Davison, B. D. 2011. Temporal dynamics of user interests in tagging systems. In *Twenty-Fifth AAAI conference on artificial intelligence*.

Zhao, J., and Ram, S. 2011. Examining the evolution of networks based on lists in twitter. In *Internet Multimedia Systems Architecture and Application (IMSAA), 2011 IEEE 5th International Conference on*, 1–8. IEEE.