Faving Reciprocity in Content Sharing Communities
A comparative analysis of Flickr and Twitter

Jong Gun Lee
LIP6-CNRS
UPMC (Paris 6)
104 Av. du President Kennedy
Paris, France (zipcode: 75016)
Tel: +33.1.4427.8877
Fax +33.1.4427.5353
jonggun.Lee@lip6.fr

Panayotis Antoniadis
LIP6-CNRS
UPMC (Paris 6)
104 Av. du President Kennedy
Paris, France (zipcode: 75016)
Tel: +33.1.4427.8877
Fax +33.1.4427.5353
panayotis.antoniadis@lip6.fr

Kavé Salamatian
Polytech-Savoie
Université de Savoie
BP 80439
Annecy, France (zipcode: 74944)
Tel: +33.4.5009.6537
Fax: +33.4.5009.6559
kave.salamatian@univ-savoie.fr

Abstract—In the Web 2.0 era, users share and discover interesting content via a network of relationships created in various social networking or content sharing sites. They can become for example *contacts*, *followers* or *friends* and express their appreciation of specific content uploaded by their peers by *faving*, *retweeting* or *liking* them depending on whether they are in flickr, twitter or facebook respectively. Then they can discover additional content of interest through the lists of favorites of their contacts and so on. This faving functionality becomes thus a central part of content sharing communities for two purposes: (a) it helps the propagation of content amongst users and (b) it stimulates users' participation and activity. In this paper, we make a first step to understand users' faving (or favoring) behavior in content sharing communities in terms of reciprocity using publicly available datasets from flickr and twitter. Do users favor content only when they really appreciate it or they often feel the need to reciprocate when their content is appreciated by one of their contacts or even by a stranger? Do people take advantage of this process to gain popularity? What is the impact of the design, the social software, of a specific community and the type of content shared? These are some of the questions that our first results help to answer.

I. Introduction

Thanks to the Web2.0 technology users share today a tremendous amount of content online, which is disseminated across social links, through collaborative rating mechanisms and sophisticated content subscription schemes. Online social networks like Facebook, MySpace, and Friendster are focusing on the social relationships between users by enforcing bidirectional social links of "friendship" through which content is diffused. Content sharing communities like YouTube, Flickr, and Twitter focus on content by allowing unidirectional links between users (e.g., contact or follower) and by providing advanced rating and filtering mechanisms (e.g., the interestingness algorithm of Flickr).

A critical functionality that is implemented in most of these sites is the action of *faving* (Flickr) or *liking* (Facebook) or *retweeting* (Twitter) a certain content item. This action itself and a complete list of favorited items are often made visible to their social network or even the whole Internet. They can also become a property of the content item itself, which could be characterized by the total number of favorites it received and gain more or less visibility depending on this value.

In content sharing communities, users can also choose to "follow" or add to their "contacts" users whose content they generally like, without this link being necessarily interpreted as a friendship bidirectional relationship. In this case they get notified immediately about their contacts' new content uploads. This way, the newly uploaded content of popular users will gain wide visibility and will have increased probability of being faved and thus further disseminated across the network. Content dissemination through contact links is further enhanced by the fact that users discover additional content of interest by visiting the lists of favorites of their contacts, the most popular of ones can thus play a role of hub for online content.

This process forms a very effective collaborative rating and filtering mechanism that enables popular content to be disseminated fast and efficiently. For instance, the available rating information can then be used to promote content of good quality in searches or include it in selected content collections provided by the corresponding online community (Flickr's explore page[1]) or by external sites (e.g., retweetradar[2]).

So, the faving and following actions in addition of communicating appreciation become this way an action of promoting content, and thus can have both psychological and practical benefits for the receiver. The main questions that we wish to answer in our research are the following:

- to what extent users participating in content sharing communities demonstrate reciprocative behavior in terms of faving content and following.
- which are the factors that can affect this behavior

The notion of reciprocity has been studied in depth in economics and game theory as a means to enforce cooperative behavior in resource sharing systems [1] but also in sociology [2] and anthropology [3]. It could be defined as a user strategy to return received favors in a similar way:

---

[1]http://www.flickr.com/explore/
[2]http://www.retweetradar.com/

responding to a positive action from others with similar a positive action and responding to a negative action with a similar negative action. Do people demonstrate such behavior in content sharing communities?

Until now reciprocity has been studied at the contact level. That is to what extent users reciprocate in the creation of following or contact links in Twitter [4] and Flickr [5] respectively, and many more popular OSNs. In this paper we extend these results by extending the notion of reciprocity to include content rating (faving) and by comparing the reciprocity behavior observed in Flickr and Twitter.

Our long-term objective is to make a step further from simple observation and understand the mechanisms that influence reciprocative behavior. Of course, there are many possible reasons for such behavior and it is not easy to distinguish between them. For example, two users might favor each other's photos simply because they really happen to like them. On the other hand, the action of faving generates satisfaction for the receiver, who in turn might feel the need to reciprocate, either to show his gratefulness or to encourage further interactions. Moreover faving behavior can be used by some to increase their own self-esteem. For example, an ambitious photographer might consider to favor the photos of many Flickr users' to attract their attention and stimulate their reciprocity behavior. The more she does this, the more her photos will become visible.

We believe that by comparing the observed behavior of users between OSNs and content sharing communities that differ in terms of social software and content shared will give us some first insights on the dependence of reciprocity behavior on the specific environment. In the future we plan to acquire similar results from the same systems after important changes in their design (e.g., the addition of favorites' list in a user's profile in Flickr) but also from similar systems (e.g., ipernity[3] which is a very close to Flickr system with a much smaller set of users mainly from Europe).

This paper is only a first step toward this direction. In the following we present and analyze results acquired from two datasets publicly available from Flickr and Twitter. More specifically, we present the per user faving activity, the level of reciprocative exchanges between pairs of users, and a more detailed analysis on the timing properties of these reciprocative exchanges.

Our results indicate that faving reciprocity does play an important role in these networks. More specifically,

1) A high percentage of favorites generated by active, Type C, users are subject to reciprocity behavior
2) The more the outgoing favorites of a user the more chances that she will receive favorites on her content
3) There is an interesting correlation between contact relationship and reciprocity in faving.
4) There are both differences and similarities between Flickr and Twitter reciprocity behavior.

The remainder of this paper is structured as follows. In

---

[3]http://www.ipernity.com

Section 2 we elaborate a little more on the differences of implementing faving activity between different social networks and content sharing communities. In Section 3 we introduce our notation and then present our analysis on reciprocity behavior of users in Flickr and in Twitter in Sections 4 and 5 respectively. In Section 6 we give a brief overview of the related work highlighting the contributions of this paper. Finally, in Section 7 we discuss the potential impact of this work and analyze the reasons that we believe that should be enriched with more measurements from different systems and at different moments in time.

II. SOCIAL SOFTWARE AND FAVORITES

Before proceeding with the presentation of our analysis it is important to get into more detail on the faving activity in social networks and content sharing communities that will help the reader that is not acquainted with these systems to understand how they operate in relation to this functionality.

There are two important design decisions on the implementation of the faving functionality:
- Visibility and privacy
- Effect on content popularity

In Flickr, favorites is an important feature of the social software and a central part of a user's home page. That is, photos having 'faved' more times have more chances to be classified as more interesting and thus appear higher in search results. The most interesting ones become part of the Flickr's home page (called 'explore'), a fact that can increase significanctly the visibility and popularity of a certain user. Interestingly, the list of a user's favorites has recently become part of a user's profile page, which lead to intense reactions from users that do not consider their favorites as part of their personal image.

Favorites play a less important role in Twitter. A favorite item of a user is visible only in her profile page but its owner is never notified about this action. However, there is another form of faving that is even more powerful than this of Flickr, which is called "retweet". A Twitter user "retweets" the posts that he likes and they automatically become part of her own stream giving them this way a more personal flavor than Flickr's favorites. However, Twitter itself, unlike Flickr, does not offer the functionality to keep a list of retweets or classify tweets based on the number of times that they got retweeted.

Note that in terms of links between users the two networks are identical in terms of functionality. Only the label changes. The term contact used by Flickr signifies a "closer" and friendly relationship between a user and its contact compared with to term following of Twitter, which has a more socially distant connotation.

In terms of privacy, a user can see who 'faved' her photos and the same holds for any person that has access to a certain photo. This is not always the case. For example, in Ipernity, a site that is very similar to Flickr and which attracted many european users from Flickr after a cencorship debate, visitors can see only the number of favorites on a certain content and not the identities of the issuers. There are additional

differences that place favorites in ipernity in a place of lower significance: For example, one needs 'two clicks' to access a user's list of favorites and this list not part of one's profile.

It is impressive that there are actually different semantics and privacy characteristics of favorites in almost every OSN or content sharing community. For example, in Facebook, unlike Flickr and Twitter, users can "like" something but they can choose to hide completely this action from their profile and there is not a centralized list of all the "liked" items of a certain user not even for her own use.

Reciprocity is a fundamental aspect of the human psychology and online behavior. So sheding some light on its characteristics and dependence on the system design choices is an important step in building content sharing systems that serve better their high level objectives. The observed differences in the current design of the faving activity in different contexts provide important evience the corresponding design decisions can affect significantly user faving behavior and how this influences content dissemination in practice.

III. NOTATION

As analyzed above, a user can create a relation between herself and another user or another content item. We will use the Flickr's terminology (favorite and contact) to refer to these two actions both for Flickr and Twitter analysis in order to facilitate the presentation of the results.[4]

We can represent these relations with a graph. The graph for contact relations will represent a user as a node and a contact relation between two users as a directed edge, if the relation is asymmetric, or undirected if the relation is symmetric. Then the in-degree of a node denotes the number of users who have added this node (user) into their contact list and the out-degree of a node is the number of users whom this node (user) has added into her own contact list.

We need a more complex graph to express user-content relations as demonstrated in Figure 1 for a case of three users. Each content item is represented by a circle inside a rectangle, the "user domain". There two types of content items: an item that is uploaded by the corresponding user is depicted with a solid line, while a dotted line signifies a pointer to a content item that is favorited by another user. This same content is represented with a solid circle in the domain of the user to whom it belongs, and the two circles are connected with a directed edge from the dotted to the solid circle. So each faving action generates a new dotted circle in the domain of the user that initiated this action and a directed edge to the corresponding dotted circle in the domain of the user that owns the corresponding content item. Figure 1 shows an example with three users, their content, and favorites as described above.

Then the in-degree of a user will correspond to the total number of incoming edges (favorites) and the out-degree,

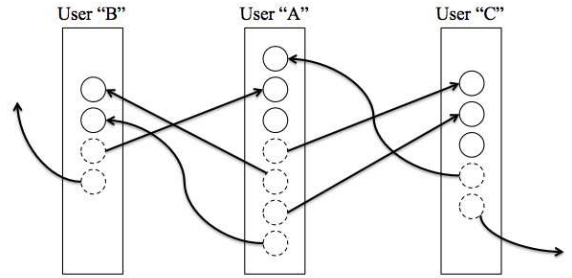Fig. 1.   Example

the total number of outgoing edges (favorites). For instance, according to Figure1, user B has in-degree of two and out-degree of four. We define formally these relations in Table I below, which provides all the required notation for both kinds of relationships, contact- and content-level, created between two users u_i and u_j.

u_i	user i
$\mathcal{U}$	all users ($u_i \in \mathcal{U}$)
Notations on favorites	
$\mathcal{F}_{ij}$	num. of u_i's favorites for u_j's content
$\mathcal{F}_{i*}$	num. of u_i's favorites for others' content, $\sum_j^{\mathcal{U}} \mathcal{F}_{ij}$
$\mathcal{F}_{*i}$	num. of others' favorites for u_i's content, $\sum_j^{\mathcal{U}} \mathcal{F}_{ji}$
$\mathcal{R}_{\mathcal{F}}(u_i)$	$\mathcal{F}_{*i}$ / $\mathcal{F}_{i*}$ (reciprocity ratio for u_i's favorites)
$\mathcal{R}_{\mathcal{F}}(u_{ij})$	$\mathcal{F}_{ji}$ / $\mathcal{F}_{ij}$ (favorite reciprocity ratio of u_i for u_j)
Notation on contacts	
$\mathcal{C}_{ij}$	1 if u_i adds u_j into u_i's contact list, otherwise 0

TABLE I
OUR TERMINOLOGY

IV. RECIPROCITY IN FLICKR

In this section, we analyze the characteristics of reciprocity with Flickr datasets. In Section IV-A we present some details of the Flickr dataset we used for our analysis. In Section IV-B, we investigate the general characteristics of each user's favorites and contacts. Then in Section IV-C, we study the reciprocity behaviors between two users.

A. Dataset

The Flickr dataset was collected for the analysis presented in [5]. This dataset contains all favorites of 789,735 Flickr users and Table II shows the brief statistics of the dataset.

Crawling duration	104 days (from Nov 2, 2006)
Num. of unique users	789,735
Num. of unique photos	11,267,320

TABLE II
BRIEF DESCRIPTION OF FLICKR DATASETS

---

[4]We refer to the action of adding a content item in one's list of favorites as 'faving', a term widely used by Flickr users today. On the other hand, when we want to refer to the action of adding a user in one's contact list we often use the Twitter's terminology 'following' which is more convenient to use.

Fig. 2. $\mathcal{F}_{*i}$ and $\mathcal{F}_{i*}$ in Flickr

Fig. 3. $\mathcal{R}_F(u_i)$ of the Users in Type C

B. Each User's Favorites

To see the characteristics of each user's favorites, we plot $\mathcal{F}_{*i}$ and $\mathcal{F}_{i*}$ of each user in Figure 2. This figure shows that about about 30% of users do not have any $\mathcal{F}_{*i}$ and about 37% of users do not have any $\mathcal{F}_{i*}$. So, we have three types of users:

- Type A, u_i who $\mathcal{F}_{*i} = 0$ and $\mathcal{F}_{i*} > 0$
- Type B, u_i who $\mathcal{F}_{*i} > 0$ and $\mathcal{F}_{i*} = 0$
- Type C, u_i who $\mathcal{F}_{*i} > 0$ and $\mathcal{F}_{i*} > 0$

Users of Type A refer to people that use Flickr mainly as a photo browsing site and they do not themselves post a significant number of photos or any at all. The second type of users are those that use Flickr only to upload their own photos, e.g., certain professional photographers. Type C users are the typical Flickr users who upload content, create social relationships in Flickr, participate in groups, etc. and they are the ones that generate the majority of the faving activity, since as depicted in Table III, the $\sum_i^{\mathcal{U}}(\mathcal{F}_{i*} + \mathcal{F}_{*i})$ of Type C users is much greater than this of the other types.

Type	# unique users	$\sum_i^{\mathcal{U}}(\mathcal{F}_{i*} + \mathcal{F}_{*i})$
Type A	243,966	5,463,886
Type B	291,798	1,285,153
Type C	253,970	39,252,466

TABLE III
NUMBERS OF USERS AND PHOTOS FOR EACH TYPE

Since our focus in this paper is on reciprocity, we will constrain ourselves in the analysis of the faving behavior of Type C users. In Figure 3 we plot $\mathcal{R}_F(u_i)$ for this type of users. By definition of $\mathcal{R}_F(u_i)$, its value is 1 when $\mathcal{F}_{*i}$ is equal to $\mathcal{F}_{i*}$. From this figure we see that about 8% of Type C users have exactly 1 of $\mathcal{R}_F(u_i)$. While $\mathcal{R}_F(u_i)$ of about 56% Type C users is less than 1, $\mathcal{R}_F(u_i)$ of about 36% of users is greater than 1, which shows that a relative small percentage of users becomes over time more 'popular' than others.

It is interesting to note that the reciprocity ratios increase with the the total number of outgoing favorites per user. Also, the majority of users, ranging from 70% for users with
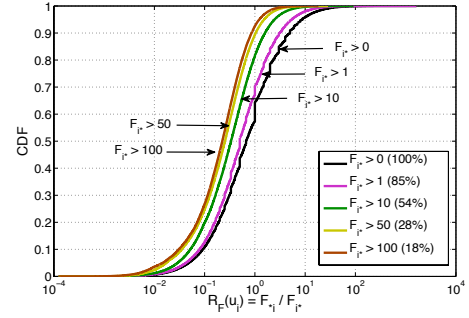
$\mathcal{F}_{i*} > 1$ to almost 85% for users with $\mathcal{F}_{i*} > 100$, have $0.1 < \mathcal{R}_F(u_i) < 10$. These results indicate that there is indeed a strong correlation between the outgoing and incoming faving activity per user. However, we need to study in more detail the user interactions in order to understand to what extent this correlation is due to reciprocative behavior. We do this in the following.

C. Reciprocity in Flickr

We select all user pairs (u_i, u_j), between whom there is at least one favorite by $\mathcal{F}_{ij} \geq 1$ or $\mathcal{F}_{ji} \geq 1$. Then we classify these pairs into two categories, non-reciprocated and reciprocated pairs. Table IV shows the number of pairs and their $\sum(\mathcal{F}_{ij} + \mathcal{F}_{ji})$. In the table, the second (third)

	num. of (i, j) pairs	$\sum (\mathcal{F}_{ij} + \mathcal{F}_{ji})$
$\mathcal{F}_{ij} > 0$ or $\mathcal{F}_{ji} > 0$	11,053,860 (100%)	27,921,971 (100%)
$\mathcal{F}_{ij} = 0$ or $\mathcal{F}_{ji} = 0$	10,046,666 (90.89%)	17,570,022 (62.93%)
$\mathcal{F}_{ij} > 0$ and $\mathcal{F}_{ji} > 0$	1,007,194 (9.11%)	10,351,949 (37.07%)

TABLE IV
FAVORITES BETWEEN TWO USERS WHO ARE CONNECTED WITH AT LEAST
ONE FAVORITE

line shows the information of non-reciprocated (reciprocated, respectively) pairs. Amongst all user pairs who made at least one favorite, the percentage of reciprocated pairs is less than 10% but $\sum(\mathcal{F}_{ij} + \mathcal{F}_{ji})$ is about 37%. In the comparison of the average favorites between two users, the average favorites of non-reciprocated couples is 1.75 but the average of reciprocated couples is much higher as 19.95. So we could say that reciprocated users play a more important role than non-reciprocated users in the number of generated favorites in the system. More specifically, the 37% of the total number of favorites of Type C users are subject to reciprocity!

We now see $\mathcal{R}_{\mathcal{F}}(u_{ij})$ for reciprocated user pairs whose $\mathcal{F}_{ij} > 0$ and $\mathcal{F}_{ji} > 0$ and we plot the CDF of $\mathcal{R}_{\mathcal{F}}(u_{ij})$ in Figure 4(a). (The inner figure shows the overall CDF distributions and the outer figure shows the cropped CDF distributions between $0.1 \leq \mathcal{R}_{\mathcal{F}}(u_{ij}) \leq 10$.) In the figure the tick straight line shows $\mathcal{R}_{\mathcal{F}}(u_{ij})$ and it implies that about 30% user pairs are equally reciprocated. This means that for a pair of users, u_i and u_j, $\mathcal{F}(u_{ij})$ is equal to $\mathcal{F}(u_{ij})$. For these

pairs which have reciprocity ratio of one, we plot their $\mathcal{F}(u_{ij})$ or $\mathcal{F}(u_{ji})$ in Figure 4(b) and we see that about 81% of them gives and takes one favorite for each other. In order to see the behavior of more reciprocated user pairs, we exclude pairs who give and take one favorite and plot $\mathcal{R}_{\mathcal{F}}(u_{ij})$ in Figure 4(c). From this figure we know that about 17% of user pairs are equally reciprocated.

One interesting question about reciprocity is to compare $\mathcal{R}_{\mathcal{F}}(u_{ij})$ with regarding to whether u_i and u_j are contacts. To answer this question in Flickr we choose two groups of user pairs, one for pairs of users who are contacts and another for pairs of users who are not contacts. In 4(a) we plot $\mathcal{R}_{\mathcal{F}}(u_{ij})$ of these two groups with the light straight and dotted lines. Interestingly, user pairs that are not contacts (the dotted line) are more reciprocated rather than user pairs that are contacts (the light straight line). However a relation between users who are not contacts could be a strict reciprocative behavior meaning that a user makes one favorite for another exactly because the other made one favorite. So we plot two groups' $\mathcal{R}_{\mathcal{F}}(u_{ij})$ after excluding the cases of $\mathcal{F}_{ij}=\mathcal{F}_{ji}=1$ in Figure 4(c). However we still observe the same behavior in this figure. This could mean that a user in Flickr may show more reciprocative behavior for users who are not in her contact list than users with whom she has a closer relationship.

Another interesting question is to study the response time of reciprocative behavior: how fast a user reciprocates a received favorite. Again, we differentiate between users that are contacts or not in order to study to what extent this relationship affects reciprocative behavior to this end. For this we extract the timestamps of the faving actions between two users and then among them we consider the timestamps that correspond to a transition of favoring direction. For example, let say $ts(k)_{ij}$ as the timestamp when k^{th} favorite is made and the direction of the favorite is $\mathcal{F}_{ij}$. Then let assume that we observe the following increasing timestamp sequence: $ts(1)_{ji}, ts(2)_{ij}, ts(3)_{ij}, ts(4)_{ji}, ts(5)_{ij}$. In this example, the reciprocating time is $ts(2)_{ij} - ts(1)_{ji}$, $ts(4)_{ji} - ts(3)_{ij}$, and $ts(5)_{ij} - ts(4)_{ji}$. For a sequence of reciprocating time between a reciprocated user pair, we calculate its mean value. We calculate mean values of all reciprocated pairs and we plot them in Figure 5. (We only plot a mean value when it is less than 100 days.) In this figure, the thick and straight line shows the CDF of reciprocating time for all reciprocated user pairs. Additionally the light and straight line shows the CDF of reciprocating time when a user pair (u_i and u_j) has a symmetric social relation ($\mathcal{C}_{ij} = \mathcal{C}_{ji} = 1$) and the dotted line implies two users in a user pair are not in a contact relationship. From this figure, we interestingly observe that when two users are not contacts, their average reciprocating time is shorter than the other case. In the case of Figure 4, this leads to the interpretation that users in Flickr tend to act more kindly and faster to other users who are not in a contact relationship with them.
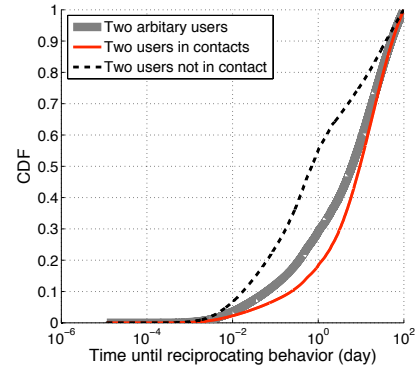

Fig. 5. Time until reciprocating behaviors

V. RECIPROCITY WITH TWITTER

A. Dataset

We use the subset of dataset used for [4]. For Tweet analysis, we use 1,686,517 tweets tagged by #iranelection. In the dataset, there are 120,319 unique users.

B. Each user's retweets

In this section we investigate Twitter users' favorite behaviors which is represented by retweets. So for each user we count $\mathcal{F}_{*i}$ and $\mathcal{F}_{i*}$ and plot them in Figure 7. This figure
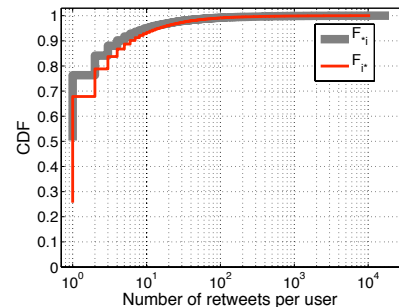

Fig. 6. Number of retweets per user ($\mathcal{F}_{*i}$ and $\mathcal{F}_{i*}$)

implies that about 50% of users do not have $\mathcal{F}_{*i}$ and about 26% of users do not have $\mathcal{F}_{i*}$. We apply the same classifying rule we used in Section IV-B and count the number of users and retweets according to types. Table V shows the statistics. This table shows that the percentage of Type A and B users

Type	# unique users	$\sum_i^{\mathcal{U}}(\mathcal{F}_{i*} + \mathcal{F}_{*i})$
All	46,402 (100%)	730,578 (100%)
Type A	23,347 (50.31%)	42,354 (5.79%)
Type B	11,879 (25.60%)	81,122 (11.10%)
Type C	11,176 (24.09%)	607,102 (83.11%)

TABLE V
NUMBERS OF USERS AND $\sum_i^{\mathcal{U}}(\mathcal{F}_{i*} + \mathcal{F}_{*i})$ FOR EACH TYPE

are about 76% but the percentage of retweeting or retweeted tweets is less than 17%. However the number of Type C users

(a) $\mathcal{R}_{\mathcal{F}}(u_{ij})$

(b) $\mathcal{F}_{ij}$ for two users whose $\mathcal{R}_{\mathcal{F}}(u_{ij})$=1

(c) $\mathcal{R}_{\mathcal{F}}(u_{ij})$ excluding $\mathcal{F}_{ij}=\mathcal{F}_{ji}$=1
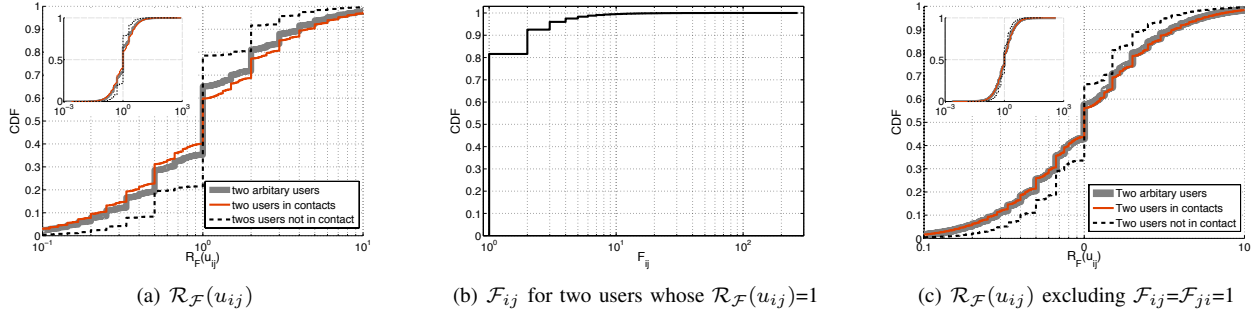
Fig. 4.   Reciprocity between Two Arbitrary Users in Flickr

is less than one-quarter, but the percentage of retweets by them is more than 83%. So Type C Twitter users are the typical Twitter users as Type C Flickr users do.

Now we look into $\mathcal{R}_F(u_i)$ of Type C users. We calculate $\mathcal{R}_F(u_i)$ of each Type C user and plot their CDF as the thick straight line in Figure 7. It shows that about 23% of Type C
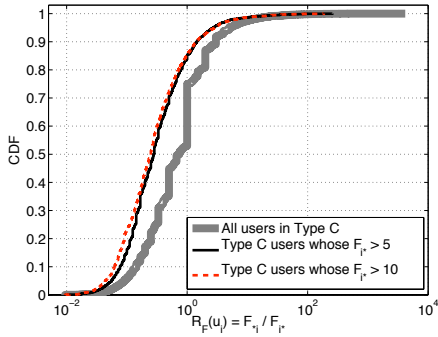
	num. of (i,j) pairs	$\sum (\mathcal{F}_{ij} + \mathcal{F}_{ji})$
$\mathcal{F}_{ij} > 0$ or $\mathcal{F}_{ji} > 0$	140,598 (100%)	365,285 (100%)
$\mathcal{F}_{ij} = 0$ or $\mathcal{F}_{ji} = 0$	135,386 (96.30%)	285,110 (78.05%)
$\mathcal{F}_{ij} > 0$ and $\mathcal{F}_{ji} > 0$	5,212 (3.70%)	80,175 (21.95%)

TABLE VI
NUMBER OF RETWEETS BETWEEN TWO USERS

Fig. 7. $\mathcal{R}_F(u_i)$ in Twitter

users have one as $\mathcal{R}_F(u_i)$, meaning that $F_{*i} = F_{*i} = 1$. To see $\mathcal{R}_F(u_i)$ of user who more actively retweet others' tweets, we additionally plot two cases, $F_{i*} \leq 5$ and $F_{i*} \leq 10$ with the light and straight line and the dotted line, respectively. These shows that these users have smaller $\mathcal{R}_F(u_i)$ than overall Type C users. In detail, about 85% of users of any case ($F_{i*} \leq 5$ or $F_{i*} \leq 10$) have less than 1 of $\mathcal{R}_F(u_i)$. Also this implies that more users make F_{i*}, their $\mathcal{R}_F(u_i)$ are closer to one, which means that they more reciprocate with others. We continue to see detail retweeting behaviors between two users in Type C in the following section.

C. Reciprocity in Twitter

In order to analyze the characteristics of the reciprocity on retweeting behaviors, we first choose all of user pairs (u_i and u_j), each of which has $\mathcal{F}_{ij} > 1$ or $\mathcal{F}_{ji} > 1$. We classify the pairs into two groups, (a) non-reciprocated user pairs and (b) reciprocated user pairs. Table VI shows their brief statistics and the second and third lines imply non-reciprocated and reciprocated user pairs, respectively. Among all user pairs, the percentage of reciprocated user pairs is only 3.7% but the percentage of retweets made by them is about 22%. For

the average retweets per user pair, non-reciprocated user pairs have 2.10 retweets but reciprocated ones have 15.38 retweets. Thus similarly to the case of reciprocated user pairs in Flickr, reciprocated user pairs in Twitter are responsible for a large percentage of retweets generated in the system. However, their importance is significantly lower in the case of Twitter.

Now we analyze $\mathcal{R}_F(u_{ij})$ for reciprocated user pairs and plot $\mathcal{R}_F(u_{ij})$ in Figure 8(a). About 25% of user pairs among all reciprocated user pairs have one as $\mathcal{R}_F(u_{ij})$ and this percentage is not much different from the percentage of user pairs who have one $\mathcal{R}_F(u_{ij})$ in Flickr. However, the percentage of equally reciprocated user pairs in Twitter is smaller than the one in Flickr. This comes from many possible reasons and one of them is that Twitter is a new media rather than a social networking service as mentioned in [4]. Notably, although Flickr is mainly a content sharing site, focusing on photos instead of news or links as in the case of Twitter, there is a stronger social dimension encouraged by the groups functionality and other social software details.

To see the numbers of retweets when $\mathcal{R}_F(u_{ij})$ is one, we plot them in Figure 8(b). About 84% of reciprocated user pairs retweet only one tweet of their partner. After excluding these retweets we plot $\mathcal{R}_F(u_{ij})$ in Figure 8(c). In this figure about 14% of user pairs are equally reciprocated. Remember that in Flickr about 17% of pairs have equal $\mathcal{R}_F(u_{ij})$. So we find that the percentages of reciprocated user pairs in Flickr and Twitter are not very different.

In Flickr, two users not in a contact relationship are more reciprocated rather than two users in contact. To see whether the respective following relationship in Twitter affects $\mathcal{R}_F(u_{ij})$, we distniguish between user pairs that have a contact (following) relationship and users pairs that do not. Then we plot two groups' $\mathcal{R}_F(u_{ij})$ in Figure 8(a). In the figure, the light and straight line shows $\mathcal{R}_F(u_{ij})$ of users in contact

and the dotted line shows the other case. Differently from Flickr, we see that contact relationships in Twitter do not affect significantly $\mathcal{R}_F(u_{ij})$. (Users in contact reciprocate only slightly more than users not in contact.) After excluding the cases where $\mathcal{R}_F(u_{ij}) = F_{ij} = F_{ji} = 1$, we find that the distributions of $\mathcal{R}_F(u_{ij})$ of two groups (users in contact or not) are almost the same. So in Twitter, a user reciprocates another's retweeting for her tweets regardless of whether or not the partner is her friend (or she follows the partner).

We investigate now the response time of reciprocative behavior for retweets in Twitter. For this we record a series of timestamps of retweets between two reciprocated users as we did with the Flickr dataset. Figure 9 shows three CDFs about reciprocating response time for two reciprocated users, (a) who are following each other, and (b) who are not following each other.
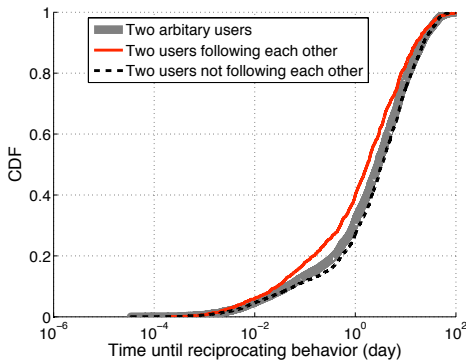

Fig. 9.   Time until reciprocating behaviors

Again, while the reponse times in Flickr are different depending on contact relationships of reciprocated users, the reciprocation time in Twitter is not significantly affected by users' contact relationhip. Unlike the reciprocated user pairs in Flickr, in Twitter reciprocated user pairs whose $C_{ij}=C_{ji}=1$ do reciprocate slightly faster.

VI. Related Work

The literature on analysis of OSNs is growing very fast touching a wide variety of disciplines (from networking to social psychology). In this section, we briefly summarize the related work that is concerned with the two main aspects addressed in this paper: reciprocity and faving activity.

- *Reciprocity*
  The concept of reciprocity plays an important role in human psychology and it has been thoroughly studied in various research areas that try to understand and influence human behavior, such as sociology[2], anthropology[3], economics[1], and politics[6]. Recently research on on-line social networking services (OSNs) has been dealing with reciprocity as a means to characterize the behavior of online users in OSNs. The main focus is given on reciprocative behavior in terms of social link creation (the contact or following relationships discussed above in the

context of Flickr and Twitter). So, the ratio of symmetric vs. asymmetric contact relationships has been studied for various OSNs, such as YouTube, ([7]), Flickr ([5], [8]), Twitter ([4], [9], [10]), Digg ([11]), and FriendFeed ([12]).
  The most related to our approach research work by Sadlon et al. [11] who study the reciprocity behavior at a content-level in the context of the Digg network. More specifically, they report strong evidence that certain users promote each other's contents in order to gain visibility in the front page of Digg which is consistent to our results.

- *Favorites of online contents*
  As we mentioned, the functionality of favorites has an important role for Web 2.0 services and it is implemented differently depending of services. Authors of [5] and [4] analyzed favorites and retweets in Flickr and Twitter, respectively without taking into account reciprocity but focusing on content difusion through contact links. Similarly, [13] investigated how voting and rating in Digg affects news popularity and [14] studied favorites in YouTube.

The major contributions of our work compared to the state of the art are the following:

- We discovered a possible correlation between users' reciprocity ratio at the content level with their contact relationship.
- We analyzed the timing properties of reciprocative behavior between pairs of users.
- We presented a comparative analysis with datasets from two popular OSN services, which can provide us with insights on the role of the system design and content on user behavior in terms of reciprocity.

VII. Discussion and future work

Faving activity in social networks and content sharing communities has many dimensions related to one's personal image, social relationships, community identity and more. We could say that it is a new type of social activity for which there is not a clear analogue in the physical world. This comes from the fact that all these actions when online can be recorded and thus become part of a user's profile and affect directly the way content is disseminated.

So, understanding the factors that affect this behavior and to what extent it can be exploited for users to become popular is both a very challenging and interesting research question.

In this paper we made only a first step toward this direction. Our results indicate that reciprocity does play an important role in the online activity of users in content sharing communities such as Flickr and Twitter, but there are significant differences.

These differences could be due to the mixture of users that participate in these different networks, the details of their design (their social software), and the type of content shared. In order to draw safer conclusions for the degree of importance of these different factors we need to explore in more depth reciprocity behavior.

(a) $\mathcal{R}_{\mathcal{F}}(u_{ij})$ (b) $\mathcal{F}_{ij}$ for two users whose $\mathcal{R}_{\mathcal{F}}(u_{ij})$=1 (c) $\mathcal{R}_{\mathcal{F}}(u_{ij})$ excluding $\mathcal{F}_{ij}$=$\mathcal{F}_{ji}$=1
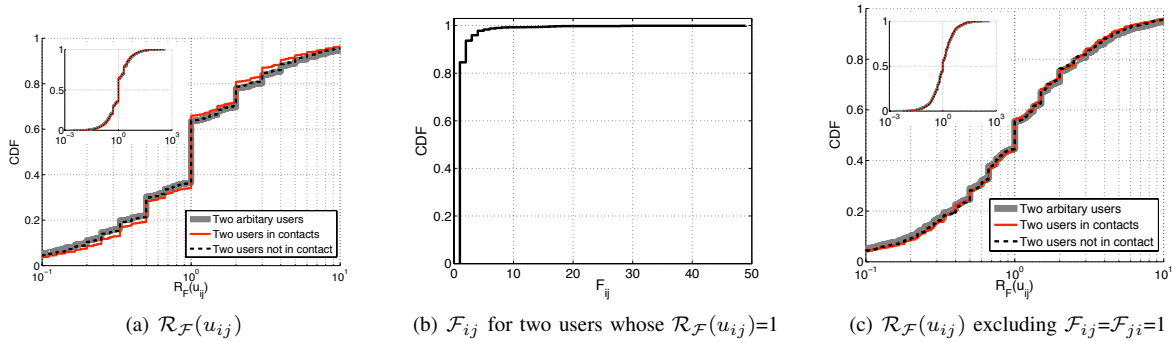
Fig. 8. Reciprocity between Two Arbitrary Users in Twitter

First, we need to study how this behavior changes over time, as new additions are made in the social software of the corresponding system. For example, did reciprocity behavior of Flickr users changed after the decision of the Flickr team to place the list of favorites in the users' profiles? A reason for this possibility could be the fact that users would be more reluctant to reciprocate favorites when they don't really like the photos of someone who have faved one of theri photos. Similarly, what was the effect of the introduction of an "official" retweeting action introduced by Twitter recently which leads to the tweet being retweeted to be placed in a user's stream with the photo of the original author? Which of the two retweet options do users prefer when their retweets are motivated by reciprocative feelings? Also, how often do users "undo" their favorites or retweets after some time? Is there a correlation with their reciprocity behavior?

Second, we need to study other similar content sharing communities. Ipernity or Tabloo for example are more or less photo sharing communities. Do we expect different reciprocity behavior in these systems? How the reciprocity behavior of twitter is compared with this of Facebook? How about with this of delicious and Digg?

Finally, we should enrich our analysis with additional information on users' activity that is easy to record and that could reveal other types of interesting correlations with their reciprocative behavior. For example, the participation of a user in groups, the volume of content generated, even demographic data could shed more light to our understanding of this interesting activity online.

To answer these questions we need more detailed datasets of a wide variety of social networks and content sharing communities over time. This is a very time consuming and computationally intensive task to be carried out by a single research institution and thus should be part of a collective effort. We believe that our first results will convince researchers to include this type of information in their future measurement analyses and thus shed more light to the factors that affect reciprocity and more generally human behavior online.

REFERENCES

[1] R. Axelrod, *The Evolution of Cooperation*. New York: Basic Books, 1984.

[2] L. Becker, *Reciprocity*. Chicago: University of Chicago Press, 1990.

[3] M. Marcel, *The Gift: Forms and Functions of Exchange in Archaic Societies [Essai sur le don]*. Glencoe, IL: The Free Press, 1954.

[4] H. Kwak, C. Lee, H. Park, and S. Moon, "What is Twitter, a Social Network or a News Media?" in *WWW'10: Proceedings of the 19th International World Wide Web Conference*, April 2010.

[5] M. Cha, A. Mislove, and K. P. Gummadi, "A measurement-driven analysis of information propagation in the flickr social network," in *In Proceedings of the 18th International World Wide Web Conference (WWW'09)*, Madrid, Spain, April 2009.

[6] P. M. Blau, *Exchange and Power in Social Life*. New York: John Wiley, 1964.

[7] J.-I. Biel, "Please, subscribe to me! Analysing the structure and dynamics of the YouTube network," 2009.

[8] C. Y. O. Nov, "Motivational, structural and tenure factors that impact online community photo sharing," *AAAI 2009*, 2009.

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

[10] B. Huberman, D. M. Romero, and F. Wu, "Social networks that matter: Twitter under the microscope," *First Monday*, vol. 14, no. 1, January 2009.

[11] E. Sadlon, Y. Sakamoto, H. J. Dever, and J. V. Nickerson, "The Karma of Digg: Reciprocity in Online Social Networks," December 2008.

[12] T. Gupta, S. Garg, A. Mahanti, N. Carlsson, and M. Arlitt, "Characterization of friendfeed - a web-based social aggregation service," March 2009.

[13] K. Lerman, "Social networks and social information filtering on digg," *CoRR*, vol. abs/cs/0612046, 2006.

[14] R. L. Santos, B. P. Rocha, C. G. Rezende, and A. A. Loureiro, "Characterizing the youtube video-sharing community," Federal University of Minas Gerais (UFMG), Belo Horizonte, Brazil, Tech. Rep., 2007.