Nepotistic Relationships in Twitter and their Impact on Rank Prestige Algorithms

Daniel Gayo-Avello[1]

Department of Computer Science, University of Oviedo, SPAIN

Abstract. Micro-blogging services such as Twitter allow anyone to publish anything, anytime. Nonetheless to say, many of the available contents can be diminished as babble or spam. However, given the number and diversity of users, some valuable pieces of information should arise from the stream of tweets. Thus, such services can develop into valuable sources of up-to-date information (the so-called real-time web) provided a way to find the most relevant/trustworthy/authoritative users is available. Hence, this makes a highly pertinent question for which graph centrality methods can provide an answer. In this paper the author offers a comprehensive survey of feasible algorithms for ranking users in social networks, he examines their vulnerabilities to linking malpractice in such networks, and suggests an objective criterion against which to compare such algorithms. Additionally, he suggests a first step towards "desensitizing" prestige algorithms against cheating by spammers and other abusive users.

Keywords. Social networks; Twitter; spamming; graph centrality; prestige.

1. Introduction

Twitter is a service which allows users to publish short text messages (tweets) which are shown to other users following the author of the message. In case the author is not protecting his tweets, they appear in the so-called public timeline and they are served as search results in response to user submitted queries. Thus, Twitter can be a source of valuable real-time information and, in fact, several major search engines are including tweets as search results.

Given that tweets are published by individual users, ranking them to find the most relevant information is a crucial matter. In fact, at the moment of this writing, Google seems to be already applying the PageRank method to rank Twitter users to that end [29]. Nevertheless, the behavior of different graph centrality methods and their vulnerabilities when confronted with the Twitter user graph, in general, and Twitter spammers in particular, are still little-known.

Thus, this paper aims to shed some light on this particular issue besides providing some recommendations for future research in the area. As it will be later discussed, user ranking in social networks cannot be an end in itself, but a tool to be used for other tasks. Hence, this author is not considering any *a priori* "good" ranking and, instead, he suggests measuring the performance of the different methods by the rankings spammers can reach with each one: the lower the spammers' rankings, the better the method.

The paper is organized as follows. First of all, a comprehensive literature review is provided. It deals with several rank prestige algorithms (some well-known and others lesser-known) which are feasible to apply to social networks; their known vulnerabilities; and some partially related work outside the scope of this study. Then, the research questions are stated and a new method to "desensitize" prestige ranking algorithms against link spamming is proposed. After that, the experimental framework in which this study was conducted is described: the dataset crawled from Twitter; the elaboration of the subset of abusive users; and the straightforward nature of the evaluation. Afterwards, results obtained with each of the different ranking methods are discussed along with the implications of the study. Finally, an in-depth analysis of the collected dataset is provided in an appendix: it provides details on the nature of the social network, the behavior of both legitimate and abusive users, in addition to some demographical analysis.

[1] Correspondence to: Daniel Gayo-Avello, Department of Computer Science (University of Oviedo) – Edificio de Ciencias, C/Calvo Sotelo s/n 33007 Oviedo (SPAIN), dani@uniovi.es

2. Literature review

A social network, despite the current association with online services, is any interconnected system whose connections are produce of social relations or interactions among persons or groups. That way, families, companies, groups of friends, or scientific production are social networks.

Social networks can be mathematically modeled as graphs and, thus, graph theory has become inextricably related to social network analysis with a long history of research. Think, for instance, of bibliometric studies that can be traced back to Lotka [20], Gross and Gross [11], Broadman [4], and Fussler [8], although the work by Garfield [9] is, with no doubt, the one with the highest impact on the daily life of nowadays scholars. However, it is not our aim to provide a survey on this topic; we recommend the reader interested in social network analysis from a Web mining perspective the corresponding chapters from the excellent books by Chakrabarti [5] and Liu [19]. Instead, for the purpose of this paper it should be enough to briefly sketch the concepts of *centrality* and *prestige*.

Both centrality and prestige are commonly employed as proxy measures for the more subtle ones of importance, authority or relevance. Thus, central actors within a social network are those which are very well connected to other actors and/or relatively close to them; this way, there exist several measures of centrality such as degree, closeness, or betweenness centrality. Centrality measures can be computed for both undirected and directed graphs; prestige, in contrast, requires to distinguish inbound from outbound connections. Thus, prestige is only applicable to directed graphs which, in turn, are the most common when analyzing social networks. As with centrality, there are several prestige measures such as indegree (the number of inbound connections, e.g. cites, in-links, or followers), proximity prestige (related to the influence domain of an actor, i.e. the number of nodes directly or indirectly linking to that actor), and rank prestige, where the prestige of a node depends on the respective prestige values of the nodes linking to it. Nonetheless to say, rank prestige is mutually reinforcing and, hence, it requires a series of iterations over the whole network.

Rank prestige is, by far, the most commonly used prestige measure and there exist a number of well-known methods to compute one or another "flavor" of such a measure. In the following subsections we will briefly review the popular PageRank, and HITS algorithms, in addition to lesser-known (although better targeted at social media) techniques such as NodeRanking, TunkRank, and TwitterRank, besides their weaknesses in different abusive scenarios.

2.1. PageRank

PageRank [27] is, in all probability, one of the best known rank prestige methods because it underlies the Google search engine [3]. The PageRank algorithm aims to determine a numerical value for each document in the Web, such a value would indicate the "relevance" or "authority" of that given document. That value, also known as PageRank, spreads from document to document following the hyperlinks –previously it must be divided by the number of outgoing links. That way, heavily linked documents tend to have larger PageRank values, and those documents receiving few links from highly relevant documents (i.e. documents with large PageRank values) also tend to have large PageRank values.

After iterating a finite (in fact a relatively short) number of steps the algorithm converges; at that moment all the nodes within the graph have got a PageRank value by means of which they can be ranked. A notable property of the algorithm is that the global amount of PageRank within the graph does not change along the iterations but it just spreads from some nodes to other ones. Thus, if the total amount of PageRank in the Web was arbitrarily fixed at 1 we could see the PageRank value for a given document as a proxy for the probability of reaching that given document by following links at random (that's why PageRank is often described as a *random surfer model*). Such a model is described by Equation (1) where $PR(p)$ is the PageRank value for webpage p, $M(p)$ is the set of webpages linking to p and $L(p)$ is the set of pages linked from p.

$$PR(p_i) = \sum_{p_j \in M(p_i)} \frac{PR(p_j)}{|L(p_j)|} \quad (1)$$

Of course, this description is an oversimplification because it makes several unreal assumptions, namely that the Web is a strongly connected graph, and that there are no sink nodes (i.e. nodes with in-links but no out-links). In order to solve this, a modified version including a *damping* or *teleportation* factor is shown in Equation (2): d is the damping probability (usually 0.15), and N is the total number of webpages in the graph.

$$PR(p_i) = (1 - d) \sum_{p_j \in M(p_i)} \frac{PR(p_j)}{|L(p_j)|} + \frac{d}{N} \quad (2)$$

2.2. HITS

Hyperlink-Induced Topic Search – HITS [15] is another algorithm to estimate the relevance of a document. The method assumes the existence of two different kinds of documents in the Web: *authorities* and *hubs*. An authority is a heavily linked document because each inbound link is a "vote" cast by the user linking that document. Conversely, a hub is a document comprising links to several authorities; therefore, hubs are valuable resources in the Web's ecosystem because they ease users the task of finding relevant information.

Because webpages can exhibit both characteristics every document in the Web has associated two different scores: namely, its authority score and its hub score. It must be noticed that HITS is not aimed to be computed across the whole Web graph but, instead, within a query dependent subgraph composed of those documents already satisfying a given query (obtained by means of a standard information retrieval system), plus those documents linked from, or linking to documents in that result set.

Therefore, HITS starts with a relatively small Web subgraph and iteratively computes both scores for every page in the graph. As it can be seen in Equations (3) and (4) authority and hub scores mutually reinforce themselves; the authority score for a given page p is the sum of the hub scores of those pages q linking to p (E is the set of edges in the graph) while the hub score for a page p is the sum of the authority weights for those pages q linked from p. It must be noticed, that with every iteration both scores must be normalized so their squares sum to 1. HITS, as PageRank, converges after a number of iterations.

Finally, although HITS was not devised to compute scores for complete graphs, but rather topic-oriented subgraphs, it can of course be applied to a whole graph and, in fact, that is the way in which we are going to apply it to the Twitter user graph.

$$auth(p) = \sum_{q:(q,p) \in E} hub(q) \quad (3)$$

$$hub(p) = \sum_{q:(p,q) \in E} auth(q) \quad (4)$$

2.3. Abusing HITS and PageRank

In spite of claims on the original PageRank paper about being *"virtually immune to manipulation by commercial interests"*, the fact is that both PageRank and HITS are prone to manipulation or, at least, they have weaknesses that can be exploited under certain circumstances.

Bharat and Henzinger [2] describe three scenarios were hyperlink analysis methods (a) can be abused, or (b) fail because of wrong assumptions. Such scenarios are: (1) mutually reinforcing relationships between hosts, (2) automatically generated links, and (3) non-relevant nodes.

The first case occurs when a document in a host is linked by many documents from a second host; because each link is counted as a single vote although they are, in all probability, published by the same author, a single individual –the one publishing the links– is earning undue importance. This phenomenon is the underlying base for the so-called *link farms* which plague the Web, and is also somewhat related to Sybil attacks[2] in reputation based systems (as one could consider Twitter).

The second scenario does not describe an abusing situation *per se*, but an assumption made by hyperlink analysis methods that eventually proved wrong: that links are published by human beings where many of them are in fact, automatically generated[3]. Although not totally equivalent, behaviors in social networks such as auto-following users, can for sure bias the results eventually obtained by algorithms such as HITS or PageRank.

The third and last scenario, namely non-relevant nodes, especially affects HITS. Bharat and Henzinger describe how documents not relevant to the query topic can drift the results if they are well connected. In contrast to the previous two scenarios, for which we can find comparable situations within a social network setting, this third one is a little more elusive. In truth, this situation can only be broadly compared with one of the most common spamming behaviors in Twitter, namely getting the more followers the better no matter the relation between the contents promoted by the spammer to the potential interests of the eventual followers.

2.4. NodeRanking

NodeRanking [28] can be considered another variation of the random surfer model with authority spreading from one node in the graph to those linked from it. The main differences between NodeRanking and PageRank are two: (1) it is devised to work on weighted graphs, and (2) the damping/teleportation parameter is not fixed for the whole graph but is computed for each node and depends on the outbound connections of the node. According to their authors, this feature makes NodeRanking *"able to adapt dynamically to graphs with different topologies."*

Thus, Equations (5) and (6) underly this algorithm. $P_{jump}(p)$ drives the damping factor for each node; as it can be seen, nodes with few outbound links have greater probability of being damped which should be interpreted as the random surfer getting bored because of the limited set of choices. $P_{choose}(p)$, in contrast, is the probability of a page p to be chosen by the random surfer; when ignoring weights in the edges it simplifies to the same assumption in PageRank, i.e. a web surfer visiting a given page q would continue to any of the p pages linked from q with equal probability.

$$P_{jump}(p) = \frac{1}{1 + |L(p)|} \tag{5}$$

$$P_{choose}(p) = \frac{1}{|L(q)|}, (q, p) \in E \tag{6}$$

2.5. TunkRank

As we have already noticed, both *PageRank* and *HITS* (in all probability the most commonly applied methods) are prone to manipulation when applied to the Web graph, in general, and to the Twitter user graph (or any other social network graph), in particular. Thus, it could be wise to propose methods tailored to the particular circumstances of social networks.

[2] A Sybil attack consists of one attacker forging several different identities which are in turn use to promote/link a given resource. The name is after Sybil Dorsett, a woman with dissociative identity disorder and the subsequent book studying her case.

[3] Think, for instance, in the role links such as *Powered by Wordpress*, or *Powered by Apache*, have played in the ranking of their respective websites.

One of such methods is the one originally proposed by Daniel Tunkelang [31] and later named *TunkRank* for obvious reasons. *TunkRank* lies on three assumptions: (1) Each user has a given *influence* which is a numerical estimator of the number of people who will read that user's tweets. (2) The attention a user pays to his followees is equally distributed; that is, if *Following(X)* is the number of followees of the user X, and X is a member of *Followers(Y)*, then the probability for X reading a Y's tweet is *1/|Following(X)|*. And (3), if X reads a tweet by Y it will retweet it with a constant probability *p*.

$$Influence(X) = \sum_{Y \in Followers\ (X)} \frac{1 + p \cdot Influence(Y)}{|Following(Y)|} \quad (7)$$

Tunkelang suggests to compute users' influence recursively and argues that, although infinite for graphs containing cycles, it would converge as powers of *p* approach zero. In fact, shortly after describing this method an implementation for *TunkRank* was publicly released[4].

Up to now no rigorous analysis of TunkRank has been performed; however, it seems plausible that, given its remarkable similarity to PageRank, it would suffer from many of the weaknesses described above (e.g. Sybil attacks, auto-following, and link spamming in general). Hence, to the best of our knowledge, this is the first thorough scholar analysis on TunkRank.

2.6. TwitterRank

TwitterRank [32] is an extension to the PageRank method which, in addition to link structure, takes into account the topical similarity between users in order to compute the influence one users wield onto the others. In that sense, TwitterRank is a topic-sensitive method which ranks users separately for different topics. Thus, in order to rank users globally (i.e. with topic independence) one should aggregate every TwitterRank value weighted according the difference topic importance within the corpus.

It must be noticed that, in addition to this, the transition probability among connected users heavily relies in both the topical similarity between users, and the number of tweets published not only by the followee, but by all the followees the follower is connected to. Certainly, these features make of TwitterRank a highly flexible method which, in theory, could easily follow topic drifts. However, we feel that such a degree of flexibility makes the algorithm difficult to scale to the number of users and tweets that are published on a daily basis[5].

Because of this, and for the sake of better comparison with the rest of rank prestige, we employed a slightly modified version of TwitterRank. The differences are the following ones: (1) instead of computing a different TwitterRank value for each user and topic to be later aggregated across topics, we aimed to compute just one TwitterRank value without relying on any topic. (2) We also changed the topical similarity measure to compare users. Instead of applying Latent Dirichlet Allocation (LDA) to find the topics, then obtain each user's distribution, and finally compute Jensen-Shannon Divergence between users' distributions, we decided to apply the much more usual cosine similarity. And lastly, (3) we simplified the way to compute the *damping/teleportation* parameter. In the original paper it was computed from the matrix of users and topics obtained by means of LDA; we, in contrast, use the ratio between the number of tweets published by a given user and the total number of tweets in the corpus.

---

[4] http://tunkrank.com

[5] The paper originally describing TwitterRank employed a dataset consisting of one million tweets published by 6,748 users mainly located in Singapore. The number of users the original authors had to rank is several orders of magnitude below the size of the dataset to be employed in this study (1.8 million). In addition to that, given the homogeneity of the user group and the size of the tweet corpus, it seems pretty clear that the number of possible topics would also be well under the number of topics that could arise in the dataset to be used in this study (27.9 million).

Equations (8) and (9) provide a description of our implementation of TwitterRank. *TR(u)* is the TwitterRank value for user *u*; γ is the probability of teleportation, a constant value between 0 and 1 for the whole graph –we used the commonly applied value of 0.15; $P(u_j,u_i)$ is the transition probability from user u_j to user u_i; $|\tau_i|$ is the number of tweets published by user u_i, and $|\tau|$ is the total number of tweets published by all the users. Lastly, $sim(u_j,u_i)$ is the similarity between users u_i and u_j which, as we have already said, was implemented as cosine similarity.

$$TR(u_i) = (1-\gamma) \sum_{u_j \in followers\ (u_i)} P(u_j,u_i) \cdot TR(u_j) + \gamma \cdot \frac{|\tau_i|}{|\tau|} \tag{8}$$

$$P(u_j,u_i) = \frac{|\tau_i|}{\sum_{a:u_j \boldsymbol{follows} u_a} |\tau_a|} sim(u_j,u_i) \tag{9}$$

Hence, TwitterRank is indeed an extension of PageRank which takes into account the topical similarity between users to weight the transitions among connected users, in addition to the number of tweets the different followees publish to establish the influence a user has on its followers.

2.7. Other related work

Apart from the previously described algorithms there have been many other approaches to inferring influence in so-called Web 2.0 environments. Most of such approaches rely not only in the user graph, but they also require additional information such as user actions (e.g. joining a group, uploading a picture, tagging a resource), or the resources and tags collected and labeled by the users in the network.

For the interested reader we recommend the works by Noll et al. [24] and Goyal et al. [10]. The first one describes the SPEAR algorithm (SPamming-resistant Expertise Analysis and Ranking) which processes data from a collaborative tagging system (e.g. del.icio.us or bibsonomy) to find the most valuable resources and users by means of a mutually reinforcement method. The second one describes different models to determine influence among users by exploiting both the social graph and the actions performed by the users within the service.

With respect to micro-blogging services such as Twitter, the work by Kwak et al. [17] is really interesting; these researchers describe a method to discover influential users in Twitter by analyzing the way in which information is diffused across the network. That is, they do not only consider if a user is following another one, but which new pieces of information s/he discovers via that followee, and how that user propagates (or not) that new information.

Nevertheless, none of these methods was considered for this study. The techniques devised by Noll et al., or Goyal et al. could (and should) be adapted to microblogging services such as Twitter but such a goal is out of the scope of this paper. With regards to the work in [17], the social graph is not the most important piece of data, the way in which information flows through it is, instead, key. Thus, for the purpose of this study, we decided to focus on those methods just relying in the social graph, namely HITS, PageRank, NodeRanking, TunkRank, and TwitterRank, in addition to a new method proposed by the author of this paper and that is introduced in the following section.

3. Research motivation

3.1. Research questions

Social networks are increasingly gaining importance in the day-to-day living of Internet users, and the contents they provide can be exploited to provide up-to-date information (the so-called real-time web). Nonetheless to say, because of the ease of publishing any content, anytime by anyone, it is ever more important to have a way to separate trustworthy/relevant/authoritative sources of information from the untrustworthy/irrelevant/un-authoritative.

Given the prior success of applying rank prestige algorithms to bring order to the Web, it seems appealing to do the same with the user graphs from social networks. Thus, the main research questions addressed in this study are the following: 1) How vulnerable to link spamming are common graph methods when applied to user graphs from social networks? And 2), is it possible to "desensitize" such algorithms in a way that makes no more necessary to detect spammers after computing the ranks, but, instead, take into account their presence and minimize their influence?

It must be said that, in addition to the graph centrality methods studied in the Literature Review this author is proposing a variation of the popular PageRank method which is much less sensitive to link abusing in social networks. Such a method, which will be thoroughly described in the following subsection, relies on a deweighting factor computed from the reciprocal links between users.

The impact of such reciprocal links (also known as nepotistic or colluding links) in the rankings have been already studied (e.g. [1, 33]); however, to the best of our knowledge most of the proposed solutions consist of first detecting the abusing subgraphs to later adapt their rankings. Our method, in contrast, incorporates that information during the ranking computation making the prior link spamming detection unnecessary.

3.2. A new method to "desensitize" prestige ranking methods against link spamming

As we have previously exposed, one of the simplest prestige measures in a network is the indegree which translated to Twitter terms is the total number of people following a user. A priori it seems a reasonable approach: the more followers a user has got the more valuable his tweets must be; otherwise, why would people follow him?

Users such as Oprah Winfrey (3.1M followers), CNN Breaking News (2.9M followers), or TIME (2M followers)[6] are almost expected to have such huge number of followers; after all, they are opinion-makers and mass media. One could even find reasons to explain the number of followers for Ashton Kutcher (4.5M), Britney Spears (4.4M), or Lady Gaga (2.8M): they are celebrities, fans are eager to know about their idols and feel they are in contact with them, etc. Which is harder to understand is how can spammers have far more followers than average users (Yardi et al. 2010; see also Section 4.2 for more details on this).

There is, however, a simple explanation for this phenomenon. As with any other social environment, Twitter has seen the emergence of its own etiquette and, for many users, following back a new follower is considered "good manners". Of course, such a behavior is not a problem *per se*, in fact it makes perfect sense: if somebody starts following you, it means (in theory) s/he is interested in what you are tweeting about; probably both of you have some common interests and, thus, it would be a good idea to follow-back that user to see what s/he is publishing.

Nonetheless to say, many users are just following back their new followers as a matter of custom and many of them are using different tools and scripts to auto-follow back their followers[7]. Once spammers took notice of this behavior it was relatively easy to get followers: spammers just needed to massively follow other users. It must be said that Twitter consider this a violation of their terms of use but spammers (and many users in general) are using this and other related methods to increase their follower count.

Hence, it seems that number of followers is not to be trusted when trying to infer a user's relevance. Indeed, different twitterers reached the conclusion that it is the follower-followee ratio which really matters. If the number of followers exceeds those of followees it seems pretty clear that the user is publishing quality content and, to the contrary, if a user is following much more people than the people is following him, it probably means that his tweets are not that interesting.

---

[6] All the follower counts in this section are as of mid February 2010.

[7] Twitter had once got an auto-follow feature that was available under request. It was later disabled because of its harmful potential.

For instance, Oprah has a follower-followee ratio of $1.67 \cdot 10^5$, CNN Breaking News of $1.04 \cdot 10^5$, TIME $2.19 \cdot 10^4$, Ashton Kutcher $1.47 \cdot 10^4$, Lady Gaga 18.08, and Britney Spears 10.43. These figures seem to make a little more sense, but what about spammers and average users? The fact is that the follower-followee ratio for spammers tend to be close to 1 and, in fact, many of them manage to have a positive ratio indeed (probably by heavily unfollowing users once they get them to follow back the spammer). With regards to common users it varies widely but, as an example, the ratio for this author is a meager 0.69.

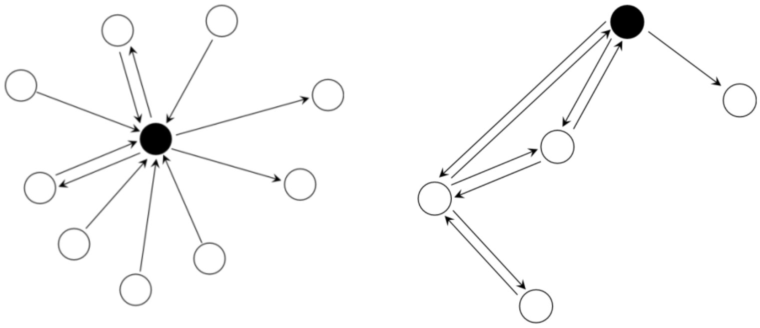

Figure 1. Three archetypical scenarios in a social network like Twitter. From left to right: an opinion-maker, a user with a close group of friends, and a user making connections by means of link exchanges (i.e. a probable spammer).

Figure 1 illustrates these three common scenarios. The first one is the typical case for a celebrity/mass-media: s/he has got lots of followers plus a few followees[8]. The second case shows the archetypical user with a close group of friends or relatives in addition to a few followees outside that group. In this scenario there are lots of connections between the users in the close group and relatively few outside. Lastly, the third scenario shows a user who has managed to get followers by mass following other users and who, in fact, has managed to have a few more followers than followees[9]. In these toy examples the follower-followee ratio is 1.75 for the heavy-followed user, 0.67 for the user with a close group of friends, and 1.14 for the presumptive spammer.

Nevertheless, let's put aside spammers for a moment so we can pay a little more attention to the follower-followee ratio. Is it just an ad hoc heuristic? Or, on the contrary, does it provide any sensible (and useful) reading? In our opinion it can be interpreted as the user's "value" regarding the introduction of new original information from the outside world into the Twitter global ecosystem. Those users which publish valuable tweets get followers who do not mind if those users are "impolite" (i.e. they do not follow back); that way they have huge number of followers but small numbers of followees and, thus, their ratios tend to be large. On the other hand, users who tend to discuss relatively personal matters with their close group of acquaintances do not get large audiences and, in turn, their ratios tend to be small (even close to zero if they follow lots of people).

But there is the spam problem. How should we tackle with it? We think the answer lies on reciprocal connections, those where two users are following each other. As we have said, many users consider this a sign of politeness but, many, especially those with huge numbers of followers, simply cannot follow-back everybody (not if they want to actually read what their followees are writing). Spammers, however, are no reading tweets and, thus, they have no constrain in the number of people to follow; especially if they aim to get a new follower in reciprocity.

In other words, reciprocal links should be under suspicion because most of the time they are used as "counterfeit currency" to increase the followers count. In consequence, we propose the *follower-followee ratio with discounted reciprocity* –see Equation (10)– which, in our opinion, captures many of the subtleties of linking in social networks.

---

[8] Just because you are Oprah or Ashton it doesn't mean you don't need to follow CNN.

[9] Probably by unfollowing users after s/he has got a follow-back link after following those users in first place.

$$ratio\_discounted = \frac{followers - reciprocal}{followees - reciprocal} \quad (10)$$

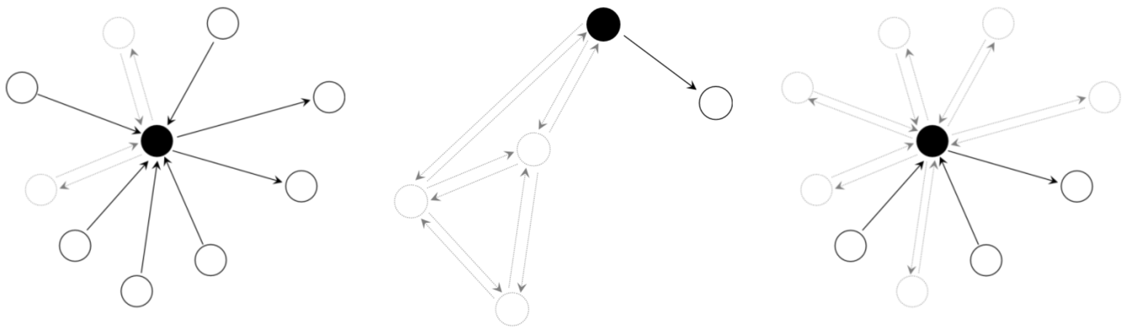

Figure 2. The same three scenarios from Figure 1 but "discounting" reciprocal connections. As it can be seen heavy-followed users are hardly affected while those users with many reciprocal connections such as small groups and spammers tend to "lose" most of their inbound links.

Figure 2 shows the same three scenarios from Figure 1 but reciprocal connections are shown in a lighter shade. Those users with many of such connections, namely spammers and common users, tend to "lose" most of their followers, while heavy-followed users are virtually unaffected. The crux of the matter is that, although for users in small close communities this could go undetected, spammers would find such a thing clearly annoying.

On the other hand, putting under suspicion all reciprocal links seems a bit obnoxious; that's why we suggest employing either the follower-followee ratio or the discounted version depending on the possible outcome: if a user would "benefit" of using the original ratio then we use the discounted one, and vice versa. Because of that, the whole name for the proposed ratio is in fact *followers to followee ratio with paradoxical discounted reciprocity* (Equation 11). For sure this can look a bit nonsensical at first, but as we are about to show it makes perfect sense.

First, let's take two users: one of them, *legit*, has 34,000 followers and 300 followees with 200 reciprocal connections; the other one, *spammer*, has 25,000 followers and 30,000 followees with 20,000 reciprocal links. The ratios for them would be 113.33 and 0.83, respectively. The ratios with discounted reciprocity would be 338 and 0.5, respectively. If the users could choose which ratio they prefer to describe themselves it seems clear that *legit* would prefer discounted reciprocity which is larger, while *spammer* would prefer the raw ratio.

However, such selfish decisions are contrary to the interpretation of the ratios. That is, *legit* would prefer assuming his reciprocal connections are not legitimate but link exchanges. On the other hand, *spammer* would prefer to count all his reciprocal connections as truly legitimate while, in all probability, this is not the case (you cannot simply follow thirty thousand people). However, both users could get us to apply the opposite ratio. For instance, *spammer* could unfollow 20 thousand people to achieve a 2.5 ratio; however, such massive unfollowing would probably make him lost most of his followers. And what about *legit*? He could massively follow most of his followers but, independently of the ratio to apply, such an action would only reduce his final value. Thus it does not seem that neither *legit* nor *spammer* would change their connections just to force a different way to compute their respective ratios.

$$paradoxical\_discounted(p) = \begin{cases} \dfrac{followers(p)}{followees(p)} & if\ followers(p) > followees(p) \\ \dfrac{followers(p) - reciprocal(p)}{followees(p) - reciprocal(p)} & otherwise \end{cases} \quad (11)$$

Lastly, the paradoxical discounted ratio is not aimed to be directly applied to users in the graph but, instead, as a weight within a given algorithm such as PageRank (see Equation 12). The only considerations when doing this is that the weight must be divided by the maximum weight found in the graph (i.e. to have the weight normalized between 0 and 1), and that the ranks must be normalized after each iteration (otherwise they would decrease towards zero).

Please notice that, by doing this, the ratio is playing the role of an externality, that is, it has an impact not on the agent responsible of its value but on third persons. In other words, a user is not affected by his own ratio but his followees (the prestige they received is de-weighted). From an economic point of view this is an appealing feature because individual spammers do not have any incentive to modify their personal behavior, although as a group –we assume they tend to weave tight networks– all of them lose out.

$$PR(p_i) = \sum_{p_j \in M(p_i)} \frac{PR(p_j)}{|L(p_j)|} \cdot \frac{paradoxical\_discounted(p_j)}{max\_paradoxical\_discounted} \tag{12}$$

4. Research design

The main goal of this study was to compare the performance of different ran prestige algorithms when applied to social networks. To attain that, a relative large dataset was needed, in addition to an objective criterion against which to compare the performance of the different methods. The dataset, described in the following subsection, was crawled by the author from Twitter. Within that dataset a subset of abusive users was selected. This was needed because the research questions underpinning this study deal with the vulnerability of centrality methods to spammers and the feasibility of "desensitizing" them to their abusive behavior. The underlying idea is to compare the different algorithms by analyzing the positioning spammers get with each of them. Thus, this section describes the dataset employed for the research, the way in which a subset of abusive users was extracted, and the simple evaluation method applied to compare the different algorithms.

4.1. Dataset description

We relied on the Twitter search API to create the dataset used in this study. To that end, we employed a query composed of frequent English stop words (e.g. *the*, *of*, *and*) in addition to forcing the results to just include tweets written in English. That query was submitted once every minute from January 26, 2009 to August 31, 2009. Nonetheless to say, the crawling was not flawless and, in fact, there were 18 days with minor network blackouts and 3 days with no tweets at all (April 26, and August 22 and 24). All in all, we collected 27.9 million English tweets corresponding to 214 days.

As we have already implied, users in Twitter can involve themselves in relationships with other users. Thus, a user can *follow* another one so that the first user can receive the tweets published by the second. This way, users can have, in Twitter parlance, *followers* and *friends*. Please notice that friends are no more than the persons a given user is following; hence we are using the term *followee* instead [13].

Using that information on followers and followees, Twitter can be represented as a directed graph. In order to build such a graph we tried to obtain followers and followees for each of the 4.98 million users appearing in the dataset. To do that, we employed the so-called *Social Graph Methods* provided by the Twitter API. This second crawl was performed after the first dataset was collected and took several months.

For the final graph we did not take into consideration links from/to users not appearing in the sample and we also dropped isolated users. In addition to this, a substantial amount of user accounts were suspended[10] at the moment of the second crawl and, hence, no information on these users' followers and followees was available. Lastly, because of the unavoidable network problems, coupled with the fact that we pushed the API a little too far, the information for a noticeable amount of users was not eventually crawled.

[10] Twitter suspends accounts when the user is violating the terms of use; suspicious use associated with spamming is, in all probability, one of the most frequent reasons.

The finally collected Twitter graph consisted of 1.8 million users; that is, 36% of the users within the original sample. It is bigger than other datasets reported in the literature (e.g. Java et al. [14] –about 90,000 users– or Choudhury et al. [6] –200,000 users) but smaller than others (e.g. Kwak et al. [16] –41.7 million users, the whole Twitter graph as of July 2009). Anyway, we think it is a fairly substantial sample given that, at the moment of collecting the dataset, the number of U.S. Twitter users was estimated between 14 [25] and 18 millions [26] and, thus, our sample is 10% that size and covers a little more than 4% of the whole Twitter graph. An in-depth analysis of this dataset is provided as an Appendix for the interested reader.

4.2. Data preparation: a subset of abusive users within Twitter

As we have described in the Research motivation section, we are interested in the feasibility of rank prestige methods less vulnerable to linking malpractice[11]. This question is highly pertinent because one or more of such methods will certainly be used to rank users to find the most relevant/trustworthy content producers. In fact, at the moment of this writing Google seems to be already applying its *PageRank* method to rank Twitter users [29] in order to offer the most relevant tweets as search results.

In order to compare different algorithms and test their respective performance we need a group of users who are actually trying to abuse Twitter linking. As many other so-called social services, Twitter is no immune to the spam problem. In addition to promoting dubious websites and products, Twitter spammers are also furiously engaged in getting followers; in fact, they have both more followees and followers than legitimate users. According to Yardi et al. (2010), they triple the number of both kinds of connections; these authors argue that "*spammers invest a lot of time following other users (and hoping other users follow them back)"*. An explanation we found highly plausible.

Whether this is done on purpose, anticipating the application of prestige algorithms to the user graph or, conversely, it is just a happy coincidence when spammers were just trying to find users to click the links they publish is irrelevant. The matter is that those spurious connections are to be treated by the eventual algorithm in the same way that legitimate links and, thus, spammers can obtain an undeserved authority within the user graph.

Thus, we decided to focus on Twitter spammers as a target group but, firstly, a method was needed to detect them. Yardi et al. (2010) describe a simple algorithm based on the presence of URLs and a selection of keywords in the tweets, in addition to the matching of certain pattern in the user names[12]. Thus, we implemented an analogous version of their technique achieving similar performance: 87.32% precision versus the 91% reported by them.

That spam detection system detected 9,369 spammers in our collection of tweets. By examining a representative sample we found that about 24% of those users were already suspended by Twitter[13], 21% of them were promoting *"money making"* techniques, about 11% were *"copywriters"* (i.e. they replicate content from other users, RSS feeds or publish plagiarized websites), 8% promote methods to get followers and/or website traffic, and the rest of them are a mixture of self-proclaimed experts in SEO, marketing, weight loss, etc.

Using a likelihood-ratio test in a way analog to that of [14] we obtained a list of distinctive terms from spammers biographies (see Table 1 for a more exhaustive list). As it was expected, terms such as *business*, *money*, *internet marketing*, *social media* and *SEO* were at the top of the list. It must be noticed that those terms are not only

---

[11] i.e. Creating links to earn reputation by exploiting knowledge about the algorithm instead of creating links as the naturally expected outcome of daily use of the service.

[12] Kwat et al. (2010) employed a different mechanism: labeling as spam those tweets including one or more *trending topics*. In addition to require additional knowledge (i.e. knowing which are the trending topics at a given time) there is also another problem with this approach: as we will show later, the average number of hashtags (usually appearing in trending topics) is well down below 3, even for spammers.

[13] Nonetheless to say, one of the main reasons for account suspension in Twitter is spamming behavior.

popular among spammers but among other Twitter users. As Yardi et al. said of them *"[they] tread a fine line between reasonable and spam-like behavior in their efforts to gather followers and attention"*. We will denote those users as *aggressive marketers* and, thus, we decided to expand the target group from pure spammers to also include those marketers.

To find them we prepared a list of terms commonly occurring in spammer biographies which were also frequently associated with marketers and other heavy-following users. By doing so, we found another 22,290 users which cannot be labeled as spammers but, we thought, could exhibit some abusive behavior with regards to connecting to other users. Table 2 shows a list of the 60 most distinctive terms for these aggressive marketers; please notice the degree of coincidence with spammers.

marketing	free	expert	online marketer
internet	help	investor	weight loss
marketer	deals	people	trump network
online	make money	network marketing	helping others
business	real estate	mlm	media marketing
money	forex	blog	marketing coach
social	coach	traffic	money making
internet marketer	home	success	help people
internet marketing	real	online marketing	forex trading
social media	news	network marketer	helping people
entrepreneur	money online	affiliate marketer	home based
affiliate	helping	making money	home business
network	tips	online business	internet entrepreneur
media	affiliate marketing	estate investor	forex trader
seo	web	small business	business coach

Table 1. The 60 most distinctive terms appearing in Twitter spammers biographies.

entrepreneur	consultant	small	owner
marketing	**social**	others	**home**
internet	**people**	**helping others**	**affiliate marketing**
real estate	**helping people**	**success**	**online marketer**
estate	**media**	**media marketing**	follow
real	**social media**	**help people**	business owner
marketer	**affiliate**	entreprenuer *(sic)*	speaker
online	marketing consultant	broker	guru
business	**web**	realtor	estate broker
internet marketing	**small business**	**estate investor**	search
helping	**coach**	**network marketing**	sem
internet marketer	**help**	estate agent	ppc
seo	**network**	agent	successful
money	**free**	**make money**	blogger
online marketing	**investor**	**expert**	**network marketer**

Table 2. The 60 most distinctive terms appearing in aggressive marketers biographies. Those terms in bold also appear in the top-60 list for spammers.

Of course, a mere similarity in biographies is not evidence of aggressive marketers trying to abuse Twitter. Because of that, we analyzed several characteristics of tweeting behavior for both spammers and aggressive marketers, and compared them to those of average Twitter users. Table 3 contains all the details and, thus, we will just summarize the most interesting findings.

First, spammers have both much more followers and followees than average users; this is consistent with the findings of Yardi et al. However, in our datasets spammers do not triple those numbers with regards to common users, but they multiply them by almost 40! Aggressive marketers are in between but much closer to spammers than to the average twitterer: they have about 15 times the number of connections than an average user.

With regards to the number of tweets published, aggressive marketers double the frequency of average users and spammers publish 7 times the number of tweets an average user does. This probably means that spammers, in contrast to marketers, are publishing tweets in an automatic fashion.

Regarding different features of the tweets themselves there are some important differences between spammers, marketers, and average users. Firstly, virtually every tweet published by a spammer contains a URL (90%); marketers use URLs in one in three tweets, while average users tend to use URLs in about one in five tweets.

Secondly, both marketers and spammers employ hashtags more than average users but the differences, although substantial, are not as pronounced as with other features. Surprisingly, the number of hashtags include by these different groups is mostly the same on average. Lastly, one feature that again highlights the robot nature of most spammers is the much lower level of retweeting, in particular, and conversations they get involved, in general. As it was expected, marketers are much more prone to retweet than average users (two times) and also get much more involved in conversations than them.

To sum up, we have described two similar, albeit not identical, groups of users –namely spammers and marketers– which exhibit several features very different from those of average users. Both have further more followers and followees than the average user, both publish more than average users, and both promote URLs more than average users. Besides, as we have stated above, the line between aggressive marketers and spammers is not always totally clear. In fact, one could assume that relatively few users will respond to the stereotypical spammer profile and, to the contrary, many users would exhibit a more or less marked spamming behavior.

Anyway, and spite of being an oversimplification, for the purpose of evaluating the different graph centrality algorithms we are going to consider a single group spammers-marketers assuming that their linking behavior, as a whole, is trying to abuse, or at least cheat, the assumptions underlying relationships in Twitter: i.e. that users should follow one another when they are genuinely interested in what that the other is publishing. Thus, in the following sections we are going to analyze in which way the global authority/reputation is divided among both groups (spammers-marketers and the rest of users), and which algorithms are less prone to be abused and, thus, provide a better ranking of users within the social network.

	Spammers	**Aggressive marketers**	**Average user**
Avg. in-degree	3203.28	1338.83	82.36
Avg. out-degree	3156.09	1245.35	82.36
Avg. # of tweets over the whole period and SD	41.25 (80.99)	12.93 (34.07)	5.60 (19.45)
% of tweets including URLs	90.42%	32.86%	18.21%
Avg. # of URLs per tweet including URLs	1.018	1.015	1.014
% of tweets including hashtags	11.54%	8.83%	7.98%
Avg. # of tags per tweet including hashtags	1.41	1.42	1.50
% of retweets over total tweets	2.97%	6.50%	2.87%
% of "conversations" over total (excluding retweets)	6.86%	21.48%	19.26%
Avg. # of users referred in conversational tweets (excluding retweets)	1.17	1.13	1.09

Table 3. Features characterizing the tweeting behavior of spammers, aggressive marketers, and average users.

4.3. Evaluation method

It must be noticed that we are not assuming any prior "correct" ranking for the users; we consider user rankings just a tool to find the most relevant source of information at a given time and, thus, a ranking algorithm will be judged by its ability to avoid abusive users achieving undeserved rankings; that's why a list of spammers and aggressive marketers was developed.

Hence, the evaluation process is quite straightforward. All of the different methods were applied to the Twitter graph to obtain a user ranking. Then, we compared the positions reached by spammers and marketers in relation to average users. The lower the rankings abusive users reach, the better the method is.

In the following section we provide the minimum, average, and median positions for the different user classes across different deciles. Such numbers will help to understand the positioning of the abusive users in relation to the rest of users in the social network. In addition to that, we will graphically show both, the percentage of abusive users found as one moves down the ranking, and the level of agreement between the different rankings.

5. Results

5.1. Prestige of abusive users when applying PageRank

About 50% of the spammers detected in the collection of tweets did not appear in the graph[14]. Those who are present account for 0.25% of the users but they agglutinate 1.4% of the total PageRank in the graph. Regarding the aggressive marketers, 98% of them appear in the graph accounting for 3.3% of the total PageRank. The acute difference from spammer to marketer presence in the graph gives an idea of the work devoted by Twitter to get rid of spammers.

Thus, the whole set of spammers and marketers which represent a mere 1.5% of the users manage to grab 4.7% of the available PageRank. With regards to their positioning in the global ranking (see Table 4, and Figures 3 and 4), 90% of spammers are, approximately, among the 60% of top ranked users and, in fact, half of the spammers appear well above the top 10% of Twitter users. On the other hand, 90% of the aggressive marketers are among the 80% of top ranked users and half of them appear above the top 20% of users.

5.2. Prestige of abusive users when applying HITS

When applying HITS to the Twitter user graph, spammers grab 5.20% of the available authority rank while aggressive marketers account for another 11%. Thus, 1.5% of the users own 16.20% of the global authority. Both, spammers and marketers are pretty good positioned (see Table 5). Half of the spammers appear at the top 10% of positions and 90% of them are among the 40% better situated users. Aggressive marketers' situation is not as good but equally impressive: half of them appear at the top 20% positions, and 90% of them are among the 60% better positioned users.

5.3. Prestige of abusive users when applying NodeRanking

When ranking users according to NodeRanking, spammers and aggressive marketers account for 1.62% and 3.86% of the global available prestige, respectively. Such amounts are about 15% larger than those achieved when using PageRank. With regards to their positioning (see Table 6), 90% of the spammers are among the 60% best situated users and half of them appear at the top 10% positions. 90% of the aggressive marketers are above 30% of the users and half of them are among the top 20% users. Such results are pretty similar to those obtained by applying standard PageRank.

5.4. Prestige of abusive users when applying TunkRank

When ranking Twitter users according to TunkRank[15], spammers and aggressive marketers account for 0.74% and 1.94% of the available global prestige, respectively. This amount is roughly half the grabbed one when applying PageRank. Attending to their positioning (see Table 7), 90% of the spammers are among the 70% of best positioned users, and half of them appear above the top 20%. Regarding aggressive marketers, there are no great differences between them and common users, although half of them are above the 40% top positioned users.

5.5. Prestige of abusive users when applying TwitterRank

When ranking Twitter users according to TwitterRank, spammers and aggressive marketers account for 0.0003% and 0.00025% of the available global prestige, respectively. In other words, using TwitterRank, both groups of abusive users reach a virtually negligible prestige (although, again, spammers manage to outperform marketers). With regards to their positioning (see Table 8), 90% of the spammers are among the top 30% users and half of

[14] Let's remember that the user graph was crawled once the collection of tweets was completed.

[15] As it was explained before, TunkRank requires a constant *p* which is the probability of users retweeting. For this experiment we employed a value of 2.87% which was found during the analysis of the dataset (see Table 3).

them appear among the 10% best positioned users. Aggressive marketers, on the other hand, seem to be slightly better positioned than average users.

The reason for these apparently contradictory results (namely, the impressive prestige reduction for spammers while still achieving top positions) is that TwitterRank distributes the prestige in a highly biased way: in fact, top 10 users account for 77% of the prestige and top 25 users for 95.5%[16]. That is, virtually all of the users in the network achieve no prestige at all and, in spite of that, spammers manage to be "one-eyed kings in the land of the blind".

5.6. Prestige of abusive users when applying PageRank with paradoxical discounting

When applying PageRank with paradoxical discounting to the Twitter user graph, spammers can only grab 0.22% of the global PageRank while aggressive marketers account for 1.05%. Thus, spammers loose -84.3%, and marketers -68.2% with regards to standard PageRank. One of the consequences of applying paradoxical discounting to PageRank is that many users reach a PageRank which is virtually zero and, hence, all of those users tie for the last position (see Table 9). 40.2% of the spammers end up in that bin while the rest of them appear among the 30% top ranked users. With regards to aggressive marketers, 55% of them reach a negligible PageRank value but the remaining 45% is among the top 30% users in the graph. This somewhat mixed results are discussed later in the paper.

5.7. Prestige of abusive users when applying PageRank to a "pruned" user graph

As it was described in a previous section, paradoxical discounting could be used to "prune" the graph which would, in turn, be ranked by means of standard PageRank. When applying this approach to the Twitter graph, spammers grab 1.84% of the available PageRank while aggressive marketers account for 4.27%. Regarding of their positioning (see Table 10), 90% of the spammers are best positioned than half of the users, and half of them are among the top 10% users. With regards to aggressive marketers, 90% of them appear among the 70% best positioned users, and half of them are among the top 20% users.

Decile	All users			Spammers			Aggressive marketers		
	Min.	Avg.	Median	Min.	Avg.	Median	Min.	Avg.	Median
9th	165,301	85,645.5	82,651	17,940.5	9,686.5	9,619	40,265	20,668.7	20,661.5
8th	330,606.5	165,308.5	165,301	32,562	17,148	17,940.5	86,683	41,880.3	40,267.5
7th	495,985.5	247,962.5	330,606.5	54,935	25,679.8	24,227.5	146,608	66,403.4	62,855
6th	661,268	330,616.5	330,606.5	81,251	36,148.6	32,562	221,907	95,327.3	86,683
5th	**826,734.5**	**413,270.5**	**413,256.5**	**117,859**	**48,788.8**	**42,085.5**	**318,219**	**129,933.9**	**115,519**
4th	992,189	495,924	495,985.5	176,590	69,949.4	54,935	446,082	171,494.7	146,608
3rd	1,156,984	578,578.1	578,550.5	282,436.5	87,641.5	66,906	612,565	222,296.9	180,888.5
2nd	1,323,166	661,232	661,268	458,004	121,866.6	81,251	846,257	284,973.9	221,907
1st	**1,487,118.5**	**743,886.1**	**743,738.5**	**846,790.5**	**178,970.1**	**98,665.5**	**1,175,085.5**	**364,625.6**	**267,299.5**

Table 4. Minimum, average, and median position for Twitter users, spammers, and marketers, across different deciles when using standard PageRank to rank the Twitter user graph.

Decile	All users			Spammers			Aggressive marketers		
	Min.	Avg.	Median	Min.	Avg.	Median	Min.	Avg.	Median
9th	165,303	82,654.5	82,656.5	4,032	2,155.7	2,215	14,107	6,602.6	6,330
8th	330,598	165,308.5	165,303	9,121	4,243.3	4,032	35,548	15,271	14,126
7th	496,058	247,962.5	247,978.5	16,608.5	6,986.2	6,131	66,727	26,847.7	23,713
6th	661,302.5	330,616.6	330,598	27,873	10,599.9	9,121	109,830.5	42,021	35,548
5th	**826,483**	**413,270.7**	**413,221**	**45,216**	**15,626.5**	**12,360**	**170,056.5**	**61,174.6**	**49,404**
4th	990,992	495,924.2	496,058	77,618.5	22,933.7	16,608.5	248,962	85,539.6	66,727
3rd	1,156,625	578,579.5	578,467.5	132,640.5	34,292.9	21,127	373,174	116,890.2	87,372
2nd	1,320,297	661,235.9	661,302.5	252,771.5	53,755.8	27,873	574,641	160,334.7	109,830.5
1st	**1,490,091**	**743,952.4**	**742,973.5**	**518,039.5**	**88,726.6**	**35,092.5**	**915,979.5**	**223,959.3**	**136,489**

Table 5. Minimum, average, and median position for Twitter users, spammers, and marketers, across different deciles when using HITS to rank the Twitter user graph.

[16] Let's compare, for instance, with PageRank which distributes only 1.2% and 2.3% of the available prestige to top 10 and top 25 users, respectively.

Decile	All users			Spammers			Aggressive marketers		
	Min.	**Avg.**	**Median**	**Min.**	**Avg.**	**Median**	**Min.**	**Avg.**	**Median**
9th	165,307.5	82,654.5	82,654.5	16,214	8,724.1	8,917	37,274.5	19,002	18,822
8th	330,602.5	165,308.5	165,307.5	30,166	15,615.7	16,214	81,564	39,015.7	37,294
7th	495,905	247,962.5	247,953.5	51,140.5	23,685.4	22,262	139,706	62,453.7	58,977
6th	661,365	330,616.5	330,602.5	75,816.5	33,459.1	30,166	212,520.5	90,346.8	81,564
5th	**826,287**	**413,270.5**	**413,309.5**	**111,405**	**45,443.1**	**39,187**	**306,308.5**	**123,786.1**	**108,975**
4th	992,190	495,924	495,905	168,385	60,910.7	51,140.5	429,445.5	163,948.5	139,706
3rd	1,175,410.5	578,578.1	578,615.5	276,355.5	83,027.3	62,142	585,996	212,707	172,703.5
2nd	1,322,298	661,232.2	661,365	448,658.5	116,781.9	75,816.5	808,880	272,550.8	212,520.5
1st	**1,487,014.5**	**743,886.2**	**743,947**	**834,146**	**172,969.8**	**93,348.5**	**1,136,582**	**349,228.7**	**256,175.5**

Table 6. Minimum, average, and median position for Twitter users, spammers, and marketers, across different deciles when using NodeRanking to rank the Twitter user graph.

Decile	All users			Spammers			Aggressive marketers		
	Min.	**Avg.**	**Median**	**Min.**	**Avg.**	**Median**	**Min.**	**Avg.**	**Median**
9th	165,313.5	82,654.5	82,651.5	29,875	15,806	15,428	70,623.5	36,099.8	36,341.5
8th	330,583.5	165,308.5	165,313.5	68,423.5	32,130.3	29,875	155,965.5	73,447	70,667.5
7th	495,982.5	247,962.5	247,988	114,073	51,371.2	48,712.5	270,273	118,997.9	110,180.5
6th	661,167.5	330,616.5	330,583.5	187,660	75,544	68,423.5	401,255	173,802	155,965.5
5th	**826,476.5**	**413,270.5**	**413,311**	**297,690.5**	**108,306.6**	**89,344.5**	**544,777**	**233,042.9**	**208,035.5**
4th	991,747.5	495,924	495,982.5	402,848.5	147,929.6	114,073	716,496	298,732.6	270,273
3rd	1,157,027.5	578,578	578,592.5	572,314	195,811.7	146,916.5	901,781.5	371,047.8	337,262
2nd	1,322,839	661,232.2	661,167.5	787,330.5	256,066.8	187,660	1,105,147	449,876.8	401,255
1st	**1,487,335**	**743,886.1**	**743,963**	**1,096,915**	**330,910.7**	**237,916.5**	**1,332,358**	**535,239.9**	**467,271.5**

Table 7. Minimum, average, and median position for Twitter users, spammers, and marketers, across different deciles when using TunkRank to rank the Twitter user graph.

Decile	All users			Spammers			Aggressive marketers		
	Min.	**Avg.**	**Median**	**Min.**	**Avg.**	**Median**	**Min.**	**Avg.**	**Median**
9th	130,344.5	83,297.9	76,499.5	24,019.5	12,713.8	12,471	130,344.5	53,960.5	52,065.5
8th	334,540.5	207,115.7	130,344.5	52,065.5	24,371.1	24,019.5	334,540.5	146,980.8	130,344.5
7th	334,540.5	249,590.7	334,540.5	76,499.5	37,128.1	38,463.5	334,540.5	209,491.2	334,540.5
6th	1,076,966	450,901.4	334,540.5	130,344.5	50,195.6	52,065.5	1,076,966	287,363.9	334,540.5
5th	**1,076,966**	**576,114.3**	**334,540.5**	**130,344.5**	**66,239.5**	**52,065.5**	**1,076,966**	**445,255.4**	**334,540.5**
4th	1,076,966	659,589.2	334,540.5	130,344.5	76,911.9	76,499.5	1,076,966	550,524.5	334,540.5
3rd	1,076,966	719,214.5	1,076,966	334,540.5	113,043.3	76,499.5	1,076,966	625,750	334,540.5
2nd	1,076,966	763,933.5	1,076,966	334,540.5	140,699.9	130,344.5	1,076,966	682,142.4	1,076,966
1st	**1,076,966**	**798,714.9**	**1,076,966**	**334,540.5**	**162,258.9**	**130,344.5**	**1,076,966**	**726,022.8**	**1,076,966**

Table 8. Minimum, average, and median position for Twitter users, spammers, and marketers, across different deciles when using TwitterRank to rank the Twitter user graph.

Decile	All users			Spammers			Aggressive marketers		
	Min.	**Avg.**	**Median**	**Min.**	**Avg.**	**Median**	**Min.**	**Avg.**	**Median**
9th	134,298	83,292.2	85,303	49,793	26,242.4	25,470	85,303	48,189	49,793
8th	312,904.5	196,352.9	134,298	134,298	52,009.5	49,793	312,904.5	92,727.9	85,303
7th	1,055,174.5	293,060	312,904.5	134,298	79,459,2	85,303	312,904.5	166,108.9	134,298
6th	1,055,174.5	483,588.6	312,904.5	312,904.5	136,279.2	134,298	312,904.5	202,816.2	312,904.5
5th	**1,055,174.5**	**597,905.8**	**312,904.5**	**312,904.5**	**171,635.4**	**134,298**	**1,055,174.5**	**298,771.7**	**312,904.5**
4th	1,055,174.5	674,116.8	1,055,174.5	1,055,174.5	197,063.9	134,298	1,055,174.5	424,819.6	312,904.5
3rd	1,055,174.5	728,553.7	1,055,174.5	1,055,174.5	319,766.9	312,904.5	1,055,174.5	514,893.8	312,904.5
2nd	1,055,174.5	171,635.4	1,055,174.5	1,055,174.5	411,591.5	312,904.5	1,055,174.5	582,417.3	312,904.5
1st	**1,055,174.5**	**197,063.9**	**1,055,174.5**	**1,055,174.5**	**483,170.8**	**312,904.5**	**1,055,174.5**	**634,959.2**	**312,904.5**

Table 9. Minimum, average, and median position for Twitter users, spammers, and marketers, across different deciles when using PageRank with paradoxical de-weighting to rank the Twitter user graph.

Decile	All users			Spammers			Aggressive marketers		
	Min.	Avg.	Median	Min.	Avg.	Median	Min.	Avg.	Median
9th	165,313.5	82,654.5	82,653.5	14,720	7,856.4	8,083	34,783	17,640.9	17,405.5
8th	330,623	165,308.5	165,313.5	27,032.5	14,024.2	14,720	76,738.5	36,527.7	34,790.5
7th	495,918.5	247,962.5	247,965	46,741	21,430.7	19,863.5	133,149	58,984.4	55,293.5
6th	661,199	330,616.5	330,623	69,424.5	30,400.2	27,032.5	203,123.5	85,750.6	76,738.5
5th	**826,346**	**413,270.5**	**413,236.5**	**102,791.5**	**41,513.2**	**35,646**	**294,846**	**117,985.8**	**103,986.5**
4th	991,809.5	495,924.1	495,918.5	155,756	55,942.5	46,741	412,074.5	156,724.4	133,149
3rd	1,156,691	578,578.1	578,616.5	251,543	76,324.3	56,803.5	564,040	203,651.2	165,030
2nd	1,322,537	661,232.3	661,199	416,371	107,509.6	69,424.5	775,863	261,087.1	203,123.5
1st	**1,486,604.5**	**743,886**	**743,962.5**	**757,974**	**159,341.3**	**85,980**	**1,097,854.5**	**334,896.5**	**244,943.5**

Table 10. Minimum, average, and median position for Twitter users, spammers, and marketers, across different deciles when using PageRank to rank a Twitter user graph "pruned" by means of paradoxical de-weighting.

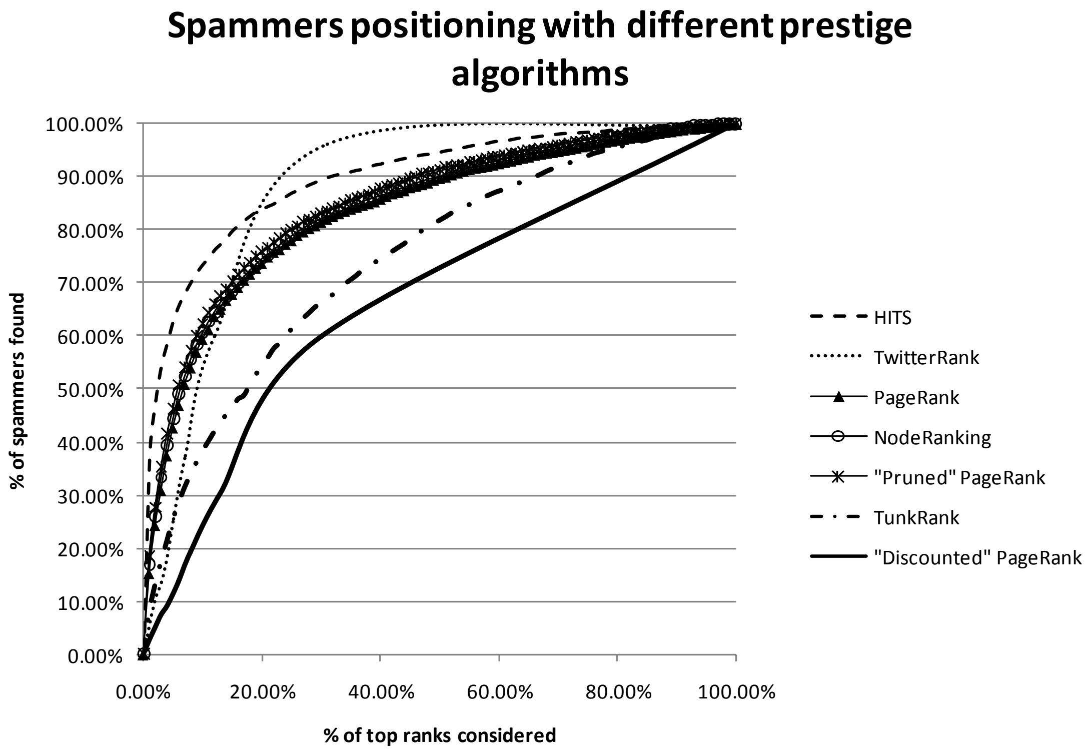

Figure 3. Percent of spammers found for different slices of the users ranking.

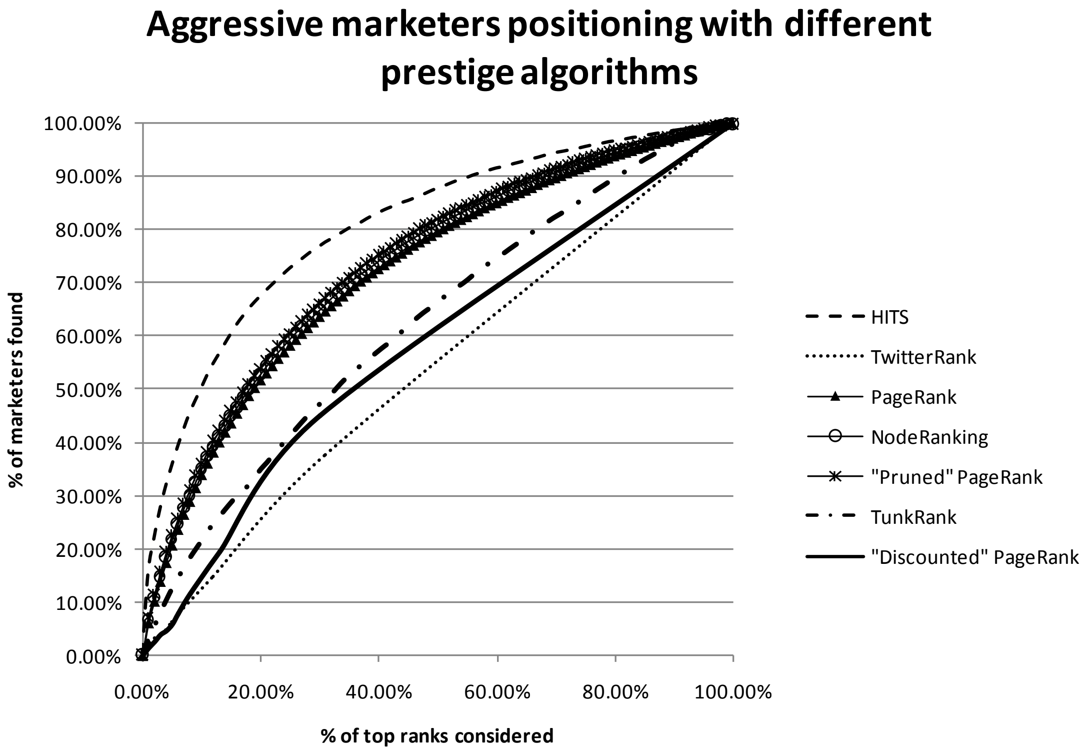

Figure 4. Percent of aggressive marketers found for different slices of the users ranking.

5.8. Agreement between the different rankings

Up to now we have shown the ability of the different algorithms to “penalize” abusive users. However, it would be interesting to check if the induced rankings are “plausible” and the level of agreement between them. Table 11 shows the top 30 users according to the different ranking algorithms.

As it can be seen, PageRank, NodeRanking, TunkRank and “pruned” PageRank exhibit a large level of agreement; all of them highly rank celebrities and personalities, news wires, and a few companies. “Discounted” PageRank, promotes several new users to the top rank, most of them musicians or related to alternative news wires.

HITS and TwitterRank are another question, HITS top rank is plagued with self-proclaimed entrepreneurs, CEOs, marketers and gurus. TwitterRank, probably because of the importance of content similarity between different users, promotes to the top of the list mainly feeds and robots interwoven in tight networks.

Table 11, however, just provides anecdotal evidence. In order to gain details on the behavior of the different ranking algorithms we compared them by means of the normalized version of Kendall distance with a zero penalty parameter ([21] and [7], respectively). Figures 5 to 7 show the agreement between the different rankings and PageRank, TunkRank, and “discounted” PageRank. In the following section we will discuss the implications of such results.

PageRank	HITS	NodeRanking	TunkRank	TwitterRank	"Discounted" PageRank	"Pruned" PageRank
aplusk (actor)	radioblogger	**aplusk (actor)**	**aplusk (actor)**	iphone_app_sale (feed)	ryanada_ms (musician)	**cnnbrk (news)**
cnnbrk (news)	brooksbayne (entrepreneur)	**cnnbrk (news)**	**cnnbrk (news)**	iphone_app_mkup (feed)	aesthetictheory (web designer)	**aplusk (actor)**
johncmayer (musician)	stephenkruiser	**johncmayer (musician)**	**stephenfry (actor)**	socialpsych (feed)	themandymoore (musician)	**johncmayer (musician)**
stephenfry (actor)	twitter_tips	**stephenfry (actor)**	**johncmayer (musician)**	psychnews (feed)	astro_127 (astronaut)	**stephenfry (actor)**
iamdiddy (musician)	bigrichb	**iamdiddy (musician)**	**theonion (news satire)**	clareelaine	barbarajwalters (news)	**iamdiddy (musician)**
jimmyfallon (comedian)	wbaustin (marketer)	**theonion (news satire)**	**jimmyfallon (comedian)**	iss_safeguard (feed)	jaygordonmdfaap (doctor)	**theonion (news satire)**
theonion (news satire)	astronautics	**jimmyfallon (comedian)**	**ryanseacrest (radio star)**	issmontserrat (feed)	**aplusk (actor)**	**jimmyfallon (comedian)**
ryanseacrest (radio star)	mattbacak (marketer)	**ryanseacrest (radio star)**	**iamdiddy (musician)**	allenwilk	**stephenfry (actor)**	**nytimes (news)**
nytimes (news)	praguebob	**mashable (news)**	**katyperry (musician)**	driveorfly	fmlteam (blog)	**mashable (news)**
mrskutcher (actress)	ann_sieg (marketer)	**nytimes (news)**	**mrskutcher (actress)**	tyneweather (robot)	**cnnbrk (news)**	**ryanseacrest (radio star)**
mashable (news)	jeanettejoy	**sarahksilverman (comedian)**	**rustyrockets (comedian)**	teesweather (robot)	**breakingnews (news)**	**sarahksilverman (comedian)**
sarahksilverman (comedian)	tmaduri	**mrskutcher (actress)**	coldplay (musician)	nyc_tweets (robot)	**johncmayer (musician)**	**mrskutcher (actress)**
rustyrockets (comedian)	joelcomm	**rustyrockets (comedian)**	**petewentz (musician)**	phoenix_tweets (robot)	**nytimes (news)**	**breakingnews (news)**
petewentz (musician)	oliver_turner	**techcrunch (news)**	**nytimes (news)**	o2apps (feed)	jimmyeatworld (musician)	**techcrunch (news)**
katyperry (musician)	andrew303	**petewentz (musician)**	**nprpolitics (news)**	apolloapps (feed)	webware (news)	**rustyrockets (comedian)**
breakingnews (news)	startuppro (entrepreneur)	**robcorddry (actor)**	danecook (comedian)	chicago_tweets (suspended)	crave (news)	**pennjillette (magician)**
techcrunch (news)	alohaarleen	**snoopdogg (musician)**	postsecret (art project)	rm_extreme (robot)	google (company)	**robcorddry (actor)**
pennjillette (magician)	orrin_woodward (guru)	**hodgman (actor)**	robdyrdek (celebrity)	thouoaksweather (robot)	imeem (music)	**timoreilly (founder of O'Reilly Media)**
robcorddry (actor)	upicks (marketer)	**ev (CEO of Twitter)**	google (company)	sb_weather (robot)	soundcloud (music)	**ev (CEO of Twitter)**
nprpolitics (news)	alicam (guru)	**nprpolitics (news)**	**pennjillette (magician)**	rm_jam (robot)	mpoppel (founder of BNO news)	astro_127 (astronaut)
snoopdogg (musician)	oudiantebi (CEO)	**breakingnews (news)**	**starbucks (company)**	isk_g1_1	Alboebno (BNO news journalist)	**hodgman (actor)**
hodgman (actor)	scotmckay (coach)	**katyperry (musician)**	chelsealately (show)	sf_tweets (robot)	felix85 (BNO news contributor)	**nprpolitics (news)**
alyankovic (musician)	clatko	**alyankovic (musician)**	**joelmchale (comedian)**	andyfranks1	Joebrooks	**petewentz (musician)**
ev (CEO of Twitter)	brat13	**mchammer (musician)**	**mashable (news)**	triciabothwell	astro_mike (astronaut)	**snoopdogg (musician)**
mchammer (musician)	robmcnealy (marketer)	**michaelianblack (comedian)**	**alyankovic (musician)**	rm_club (robot)	jeffbarr (Amazon evangelist)	**katyperry (musician)**
michaelianblack (comedian)	0boy	**timoreilly (founder of O'Reilly Media)**	ichcheezburger (humor)	rm_harder (robot)	hollymadison123 (model)	**mchammer (musician)**
jon_favreau (actor)	seanmalarkey (entrepreneur)	**pennjillette (magician)**	alancarr (comedian)	isk_g1_10 (robot)	sonsofnero (designer)	**alyankovic (musician)**
joelmchale (comedian)	suburbview	**jon_favreau (actor)**	jason_mraz (musician)	isk_g1_17 (robot)	twitterapi (Twitter API)	**jon_favreau (actor)**
starbucks (company)	coolsi (marketer)	**starbucks (company)**	marthastewart (entrepreneur)	isk_g1_18 (robot)	warped09 (music festival)	jack (Co-founder of Twitter)
timoreilly (founder of O'Reilly Media)	caseywright (CEO)	google (company)	markhoppus (musician)	isk_g1_16 (robot)	katehavnevik (musician)	**starbucks (company)**

Table 11. Top-30 users according to different ranking algorithms. A brief description is provided with the user alias. Those users shown in bold appear in the PageRank list. Those shaded appear at least in another list.

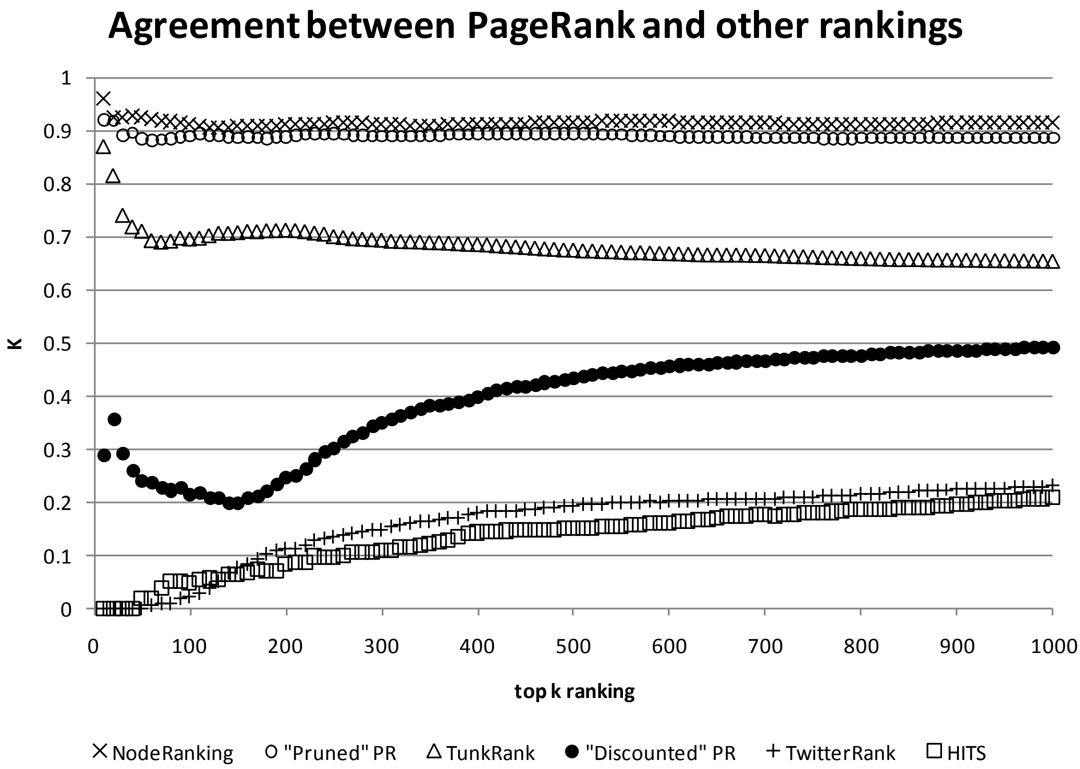

Figure 5. Aggrement between PageRank and the rest of rankings.

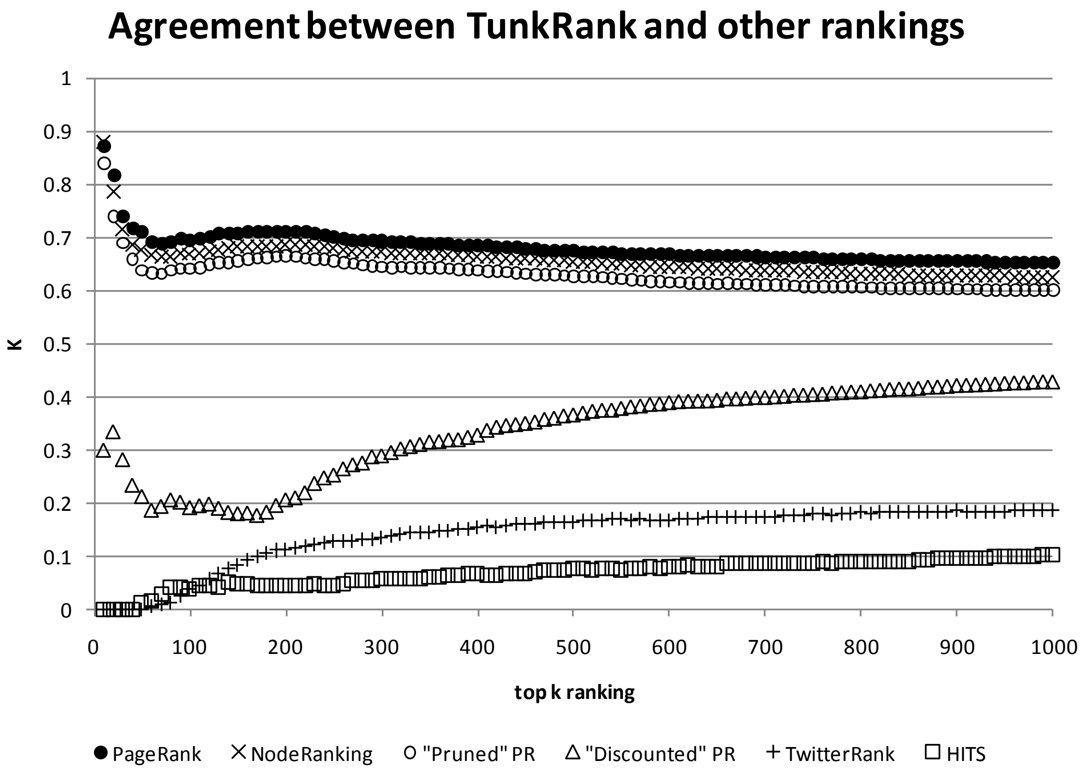

Figure 6. Agreement between TunkRank and the rest of rankings.

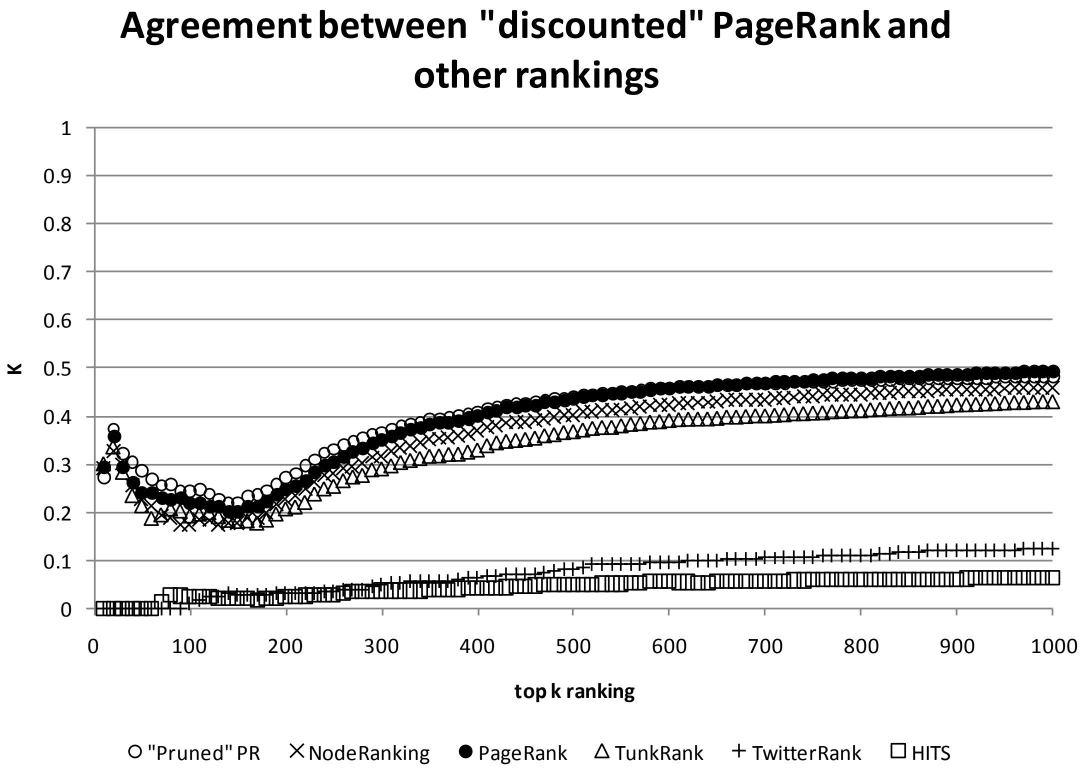

Figure 7. Agreement between PageRank with paradoxical discounting and the rest of rankings.

6. Discussion of results

As we have explained before, our approach to evaluate the available ranking algorithms in the context of social networks was not based on an *a priori* "good" ranking but, instead, their respective robustness to "gaming" by abusive users, i.e. their "ability" to penalize spammers and aggressive marketers who try to reach better positions by exchanging links instead of providing better content.

Because of the common knowledge about PageRank and, additionally, the fact that it seems to be applied by Google to rank Twitter users in their real-time web search engine, we decided to take that method as the baseline against which the rest of techniques should be compared.

The analysis of the results obtained by PageRank when applied to the Twitter user graph support our initial concern, that is, one user's PageRank is not only a measure of his value within the Twitter ecosystem, but also a consequence of the "tips and tricks" one can employ when establishing relationships within the social network. This is the only plausible explanation for spammers being much better positioned than aggressive marketers when the value of the contents they provide is virtually negligible.

There are two methods which are extremely similar to PageRank both in terms of ranking abusive users, and in terms of ranking aggrement: NodeRanking and "pruned" PageRank. With respect to NodeRanking, the similarities are unsurprising given that both NR and PR are highly related. Perhaps NodeRanking could reach its full potential with a weighted Twitter graph; however, the way in which such a weighted graph could be inferred from Twitter data (e.g. taking into account the number of mentions or retweets among users, or their content similarity) is left for future research.

The similarity between the results obtained by PageRank and PageRank applied to the "pruned" Twitter graph are somewhat expected; however, they deserve a deeper analysis because they support another point of this author. Let's remember that the "pruned" graph was obtained by removing those users (and their in- and out-links) with zero de-weighting which, in turn, was computed taking into account reciprocal links between users.

One of the arguments of this author is that discounting reciprocal links is a pretty fine way to separate users contributing to the ecosystem, from those with little or no value at all. The results obtained with the total Twitter graph and the "pruned" graph are virtually the same and, thus, we'll take that as supportive of the goodness of our proposed method.

In contrast, there are two methods which greatly differ not only from PageRank but also from the other techniques, namely HITS and TwitterRank. Each of them exhibits different problems when applied to the Twitter graph.

HITS underperforms PageRank with respect to both spammers and marketers, and the induced ranking is very different from the other rankings. As it can be seen in Table 11 the top list produced by HITS is plagued by mostly irrelevant users (at least compared with the top lists produced by the other methods). If we checked the number of followers and followees for those users we could see that the ratio for most of them is close to 1 and, in fact, most of them have got a large number of reciprocal links. In fact, because of the very nature of HITS, this algorithm is virtually inoperative when confronted with a relatively small number of users weaving a tight network of reciprocal connections. Let's remember that many spammers and marketers tend to massive follow other users in order to gain a follow-back link. Because of this, when computing hub scores those users' values tend to grow very fast; then, those hub scores are used to compute authority scores for their followees (which are mostly spammers and following them back). It is clear that after several iterations those users with lots of reciprocal links earn an undeserved amount of authority. Hence, the HITS algorithm is not advisable at all to rank users within social networks without previously "cleaning" the graph.

The results achieved by TwitterRank were very disappointing. Conceptually, it is a very appealing method: it provides ways to incorporate both content similarity measures and transition probabilities into the ranking. Some way, however, these appealing ideas seem to fail: as it can be seen from the top list, most of the users are feeds and robots, many of them highly related (even with strikingly similar names). To be fair it must be said that modifying a topic-sensitive method to operate globally is, perhaps, pushing too hard the technique. However, given that even the simplified version implemented for this research (using cosine similarity instead of LSA) is (1) much more computationally expensive than the rest of methods surveyed, and (2) it requires much more data (namely, the tweets) to obtain the rankings, it seems not at all recommendable, especially when other available methods (e.g. TunkRank) are faster and provide much better results (at least when applied globally to the complete user graph).

Lastly, there are one method clearly outperforming PageRank with respect to penalization of abusive users while still inducing plausible rankings: TunkRank. It is certainly similar to PageRank but it makes a much better job when confronted with "cheating": aggressive marketers are almost indistinguishable from common users –which is, of course, desirable; and spammers just manage to grab a much smaller amount of the global available prestige and reach lower positions –although they still manage to be better positioned than average users. In addition to that, the ranking induced by TunkRank certainly agrees with that of PageRank, specially at the very top of the list, meaning that many users achieving good positions with PageRank should also get good positions with TunkRank. Thus, TunkRank is a highly recommendable ranking method to apply to social networks: it is simple, it induces plausible rankings, and severely penalizes spammers when compared to PageRank.

With regards to the method devised by this author, "discounted" PageRank, the results are not conclusive. It seems to outperform PageRank –and even TunkRank– because the amount of prestige grabbed by abusive users is smaller and their rankings lower than when applying standard PageRank. Nevertheless, it has two issues which deserve further research.

On one hand, the induced ranking could be labeled as "elitist" because most of the users –about 70%– tie for the last position. It could be argued that this is unsurprising given that 16% of the users in the graph have got a zero de-weighting factor what we interpret as their contributions being "worthless" for the network as a whole. Actually, such results are consistent with the well-known participation inequality [23], and with a recent study revealing that 75% of the users just publish a tweet every 9 days, and 25% of the users do not tweet at all [12]. Hence, this could be considered a minor issue.

On the other hand, "discounted" PageRank exhibits a fairly distinctive curve (see Figure 7(when comparing its agreement with other rankings –obviating the underperforming HITS and TwitterRank. The agreement is much lower than, for instance, that found between PageRank and TunkRank, but the most striking behavior is the local maximum at the top positions, followed by a relatively large trough, to eventually stabilize. We found several lesser-known users at top ranks and, after studying them, we concluded that most of them have one or more "famous" followers who, in many cases, they manage to outrank. We have denoted this as the "giant shoulders" effect and it explains not only the trough at the head of the list but the smaller agreement for the rest of the ranking: many of the top users from PageRank or TunkRank are a little behind of lesser-known users they are following. This is aesthetically displeasing, at least, and the effect it can exert in the applications of the ranking is still to be explored. Nevertheless, tackling with this and the former issue is left for future research.

7. Implications, conclusions, and future work

This study makes four main contributions. First, when applying graph prestige algorithms to social networks, ranking is not only a measure of a user's value but also the result of "gaming" the algorithm by means of relationship links. The fact that spammers –who contribute no valuable content– are consistently better positioned than marketers –who contribute somewhat valuable information– no matter the method employed supports this assert.

Second, evaluating ranking should not be a point in itself; it should, instead, be evaluated within an objective context. Avoiding abusive users to reach undeserved rankings is a good metric to compare the performance of different algorithms.

Third, TunkRank is an obvious candidate to rank users in social networks. Although highly related to PageRank, TunkRank outperforms it with respect to penalizing abusive users while still inducing plausible rankings. In addition to that, it is simple to implement and computationally cheap –at least as cheap as PageRank.

And fourth, de-weighting the influence of a user by discounting reciprocal links seems to be a good way to separate those users contributing valuable contents to the global ecosystem, from those with little to no value at all. This is supported by the fact that when applying PageRank to both the complete version of the Twitter graph and to a "pruned" version where users with zero de-weighting were removed we obtained virtually the same results.

This study opens several lines of research. First, the rankings induced by the different methods should be analyzed in other contexts, for instance, as a way to rank content providers in order to find relevant information within a social network. Second, TunkRank is not immune to manipulation and, thus, its vulnerabilities should be thoroughly studied (e.g. Sybil attacks could be a starting point). And third, a deeper analysis of the role of nepotistic links, in general, and the "discounted ratio" described in this paper, in particular, is needed.

8. Acknowledgements

The author would like to thank David J. Brenes, Miguel Fernández, Fernando Zapico, and Diego Guerra for their help during the Twitter dataset collection. He is also in debt with Miguel Fernández for comments on an early draft of this paper. This work was partially financed by grant UNOV-09-RENOV-MB-2 from the University of Oviedo.

9. References

[1] R. Baeza-Yates, C. Castillo, and V. López. "PageRank increase under different collusion topologies," Proceedings of the 1st International Workshop on Adversarial Information Retrieval on the Web (AIRWeb), 2005, pp. 17—24.

[2] K. Bharat, and M. Henzinger. "Improved algorithms for topic distillation in hyper-linked environments," Proceedings of the 21st International ACM SIGIR Conference, 1998, pp. 104—111.

[3] S. Brin, and L. Page. "The Anatomy of a Large-Scale Hypertextual Web Search Engine," Computer Networks and ISDN Systems, vol. 30, no. 1-7, 1998, pp. 107—117.

[4] E. Broadman, “Choosing physiology journals,” Bulletin of the Medical Library Association, vol. 32, no. 4, 1944, pp. 479—483.
[5] S. Chakrabarti. Mining the Web – Discovering Knowledge from Hypertext Data, Morgan-Kauffman, 2002.
[6] M.D. Choudhury, H. Sundaram, A. John, D.D. Seligmann, and A. Kelliher. Birds of a Feather: Does Attribute Homophily Impact Information Diffusion on Social Media? Under review.
[7] R. Fagin, R. Kumar, D. Sivakumar. “Comparing top k lists,” Proceedings of the 14th anual ACM-SIAM symposium on Discrete Algorithms, 2003, pp. 28—36.
[8] H.H. Fussler. “Characteristics of the research literature used by chemists and physicists in the United States,” 1949, In: (A.J. Meadows ed.) The origins of information science, Taylor Graham Publishing, 1987.
[9] E. Garfield. “Citation Analysis as a Tool in Journal Evaluation,” Science, vol. 178, 1972, pp. 471—479.
[10] A. Goyal, F. Bonchi, and L.V.S. Lakshmanan. “Learning Influence Probabilities in Social Networks,” Proceedings of the Third ACM International Conference on Web Search and Data Mining, 2010, pp. 241—250.
[11] P.L.K. Gross, and E.M. Gross. “College libraries and chemical education”, 1927, In: (A.J. Meadows ed.) The origins of information science, Taylor Graham Publishing, 1987.
[12] B. Heil, and M. Piskorski, M. New Twitter Research: Men Follow Men and Nobody Tweets, 2009. Available at: http://blogs.hbr.org/cs/2009/06/new_twitter_research_men_follo.html (accessed 29 March 2010)
[13] B.A. Huberman, D.M. Romero, and F. Wu. “Social networks that matter: Twitter under the microscope,” First Monday, vol. 14, no. 1—5, 2009. Available at: http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2317/2063 (accessed 29 March 2010)
[14] A. Java et al. “Why We Twitter: Understanding Microblogging Usage and Communities,” Proceedings of the 9th WebKDD and 1st SNA-KDD workshop on Web Mining and Social Network Analysis, 2007, pp. 56—65.
[15] J.M. Kleinberg, J.M. “Authoritative sources in a hyperlinked environment”, Proceedings of the ninth annual ACM-SIAM symposium on Discrete algorithms, 1998, pp. 668—677.
[16] H. Kwak et al. “What is Twitter, a Social Network or a News Media?,” Proceedings of the 19th World Wide Web Conference, 2010. Available at: http://an.kaist.ac.kr/~haewoon/papers/2010-www-twitter.pdf (accessed 29 March 2010)
[17] H. Kwak et al. “Finding Influentials from the Order of Information Adoption and Effective Readers in Twitter,” Proceedings of the International AAAI Conference on Weblogs and Social Media, 2010.
[18] S.D. Levitt, and S.J. Dubner. “Would a Roshanda by Any Other Name Smell as Sweet?,” Freakonomics, William Morrow, 2005.
[19] B. Liu. Web Data Mining – Exploring Hyperlinks, Contents, and Usage Data, Springer, 2006.
[20] A.J. Lotka. “The frequency distribution of scientific productivity,” 1927, In: (A.J. Meadows ed.) The origins of information science, Taylor Graham Publishing, 1987.
[21] F. McCown, and M.L. Nelson. “Agreeing to Disagree: Search Engines and Their Public Interfaces,” Proceedings of the 7th ACM/IEEE-CS joint conference on Digital Libraries, 2007, pp. 309—318.
[22] A. Mislove et al. “You Are Who You Know: Inferring User Profiles in Online Social Networks,” Proceedings of the Third ACM International Conference on Web Search and Data Mining, 2010, pp. 251—260.
[23] J. Nielsen. Participation Inequality: Encouraging More Users to Contribute, 2006. Available at: http://www.useit.com/alertbox/participation_inequality.html (accessed 29 March 2010)
[24] M.G. Noll et al. “Telling Experts from Spammers: Expertise Ranking in Folksonomies,” Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval, 2009, pp. 612—619.
[25] N. O’Neill. Twitter Roars Past 14 Million U.S. Users, 2009. Available at: http://www.socialtimes.com/2009/04/twitter-14-million/ (accessed 29 March 2010)
[26] A. Ostrow. 18 Million Twitter Users by End of 2009, 2009. Available at: http://mashable.com/2009/09/14/twitter-2009-stats/ (accessed 29 March 2010)
[27] L. Page et al. The PageRank Citation Ranking: Bringing Order to the Web, 1998, Available at: http://dbpubs.stanford.edu/pub/1999-66 (accessed 29 March 2010)

[28] J.M. Pujol, R. Sangüesa, and J. Delgado. "Extracting Reputation in Multi Agent Systems by Means of Social Network Topology," Proceedings of the first international joint conference on Autonomous agents and multiagent systems, 2002, pp. 467—474.
[29] D. Talbot. How Google Rank Tweets, 2010. Available at: http://www.technologyreview.com/web/24353/page1/ (accessed 29 March 2010)
[30] S.L. Tauro, C. Palmer, G. Siganos, M. Faloutsos. "A Simple Conceptual Model for the Internet Topology," Proceedings of IEEE Global Telecommunications Conference, 2001.
[31] D. Tunkelang. A Twitter Analog to PageRank, 2009. Available at: http://thenoisychannel.com/2009/01/13/a-twitter-analog-to-pagerank/ (accessed 29 March 2010)
[32] J. Weng et al. "TwitterRank: Finding Topic-sensitive Influential Twitterers," Proceedings of the third ACM international conference on Web Search and Data Mining, 2010, pp. 261—270.
[33] H. Zhang et al. "Making Eigenvector-Based Reputation Systems Robust to Collusion," Proceedings of the third workshop on Algorithms and Models for the Web Graph, 2004, pp. 92—104.

Appendix. In-depth Analysis of the Twitter Dataset

The study described in this paper relied on a Twitter dataset collected by the author along 2009. The dataset is composed of two different parts: a collection of 27.9 million English tweets, and a user graph comprising 1.8 million users and 134 million connections. In this Appendix we provide an in-depth analysis of such dataset: we describe both the network characteristics and several demographical features of the users in the network.

Table A-1 shows some statistics describing the collected graph and compares it with the graph previously built by Java et al. [14], and with the whole Twitter crawl by Kwak et al. [16]. Some of the values for those graphs are fairly similar –or at least comparable– while those with bigger differences can be attributed, in all probability, to the Twitter growth in the last two years, in addition to sampling artifacts. In fact, when comparing the graphs by this author and Java et al., the increase in the average degree, the size of both the largest WCC and SCC, and the clustering coefficient is consistent with a growth in the number of users together with a larger number of connections between them.

Property	**Twitter sample (2009)**	**Twitter graph (2009) [16]**	**Twitter in 2007 [14]**
Total nodes	1,804,131	41.7M	87,897
Total links	134,500,669	1,047M	829,247
Average Degree	74.55	25.11	18.86
Indegree Slope	-1.33	-2.276	-2.4
Outdegree Slope	-1.516	N/A	-2.4
Degree Correlation	0.490	N/A	0.59
Diameter	6	4.8 (effective diameter [30])	6
Largest WCC size	1,800,132 (99.78%)	N/A	81,769 (93.03%)
Largest SCC size	1,688,395 (93.58%)	N/A	42,900 (48.81%)
Clustering coefficient	0.151	N/A	0.106
Reciprocity	0.48	N/A	0.58

Table A-1. Network properties for three different Twitter graph crawls. From left to right, the graph collected by this author, a whole complete graph crawled in 2009 [16], and a subgraph collected in 2007 [14].

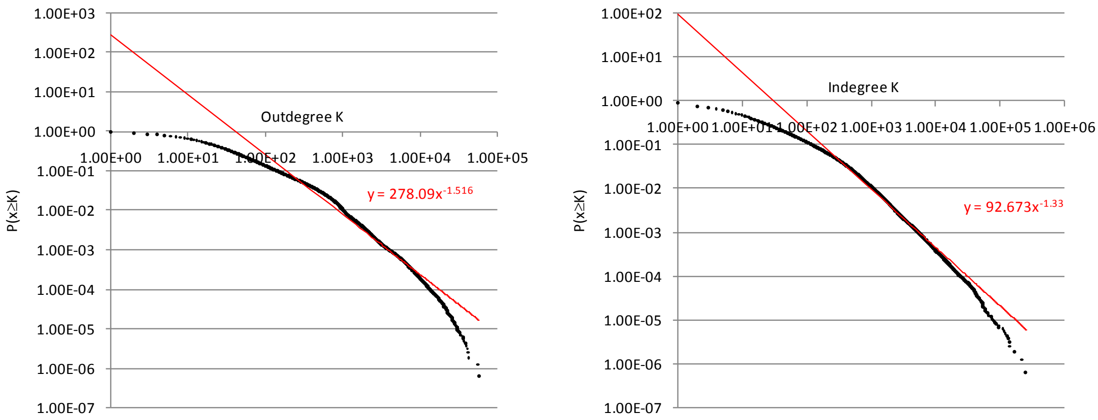

Figure A-1. Outdegree and indegree distributions in the Twitter graph. Both exhibit a power law exponent (-1.516 for the outdegree and -1.33 for the indegree). Surprisingly, Kwak et al. argue that the follower distribution for the whole Twitter graph does not follow a power law [16].

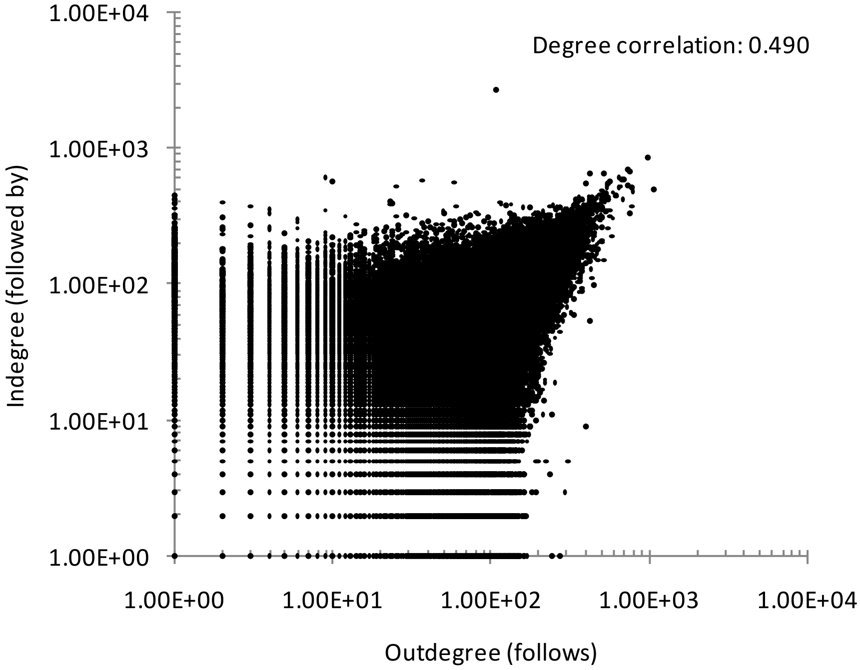

Figure A-2. Scatter plot showing the correlation between in- and out-links in the Twitter user graph. The 0.49 correlation is similar, although somewhat smaller than the reported in [14].

During the graph collection profile information was obtained for each user, namely the user's name, location, biography and website (see Table A-2 for an extract of such profiles). Such information was employed to determine the geographical location, gender, and age of the users.

Bambaloo	Emma Bullen	London		
becky_mallery	becky mallery	Kent (but second home london!)	MSc Positive Psychology Student, Author and Coach	
Burkjackson	burkjackson	iPhone: 46.181351,-123.818344	creative, thinker, father, dreamer...	http://www.Jackson5Home.com/
natekoechley	Nate Koechley	Home	I split time between online & San Francisco. One wife, one daughter, two cats, three vices & four eyes. Now: Outspark's VP of UX. Previously Yahoo! & YUI.	http://nate.koechley.com/blog
pete_watson	pete_watson	England - home of the uprising		
scottisafool	Scott Lovegrove	50.69052061,-1.93774726	Software tester for a telecomms company in South England	http://scottisafool.spaces.live.com/
ShaLaylaj	ShaLayla J. Simmons	Too far from home	I've actually been hailed as probably the most significant woman who's ever existed (I was hoping for Jet Beauty of the Week, but I can make this work).	
Swpatrick	Patrick Littlemore	London, South & West	London Career Estate Agent (Lettings). Married and father of one amazing angel. An Australian who loves London!	http://tiny.cc/Tenants
TheRealSani	Sandra S.	Home, sweet home	Hi, I'm Sani and i believe in GOD. I love peace, butterflys, Retro, Mick Jagger and how my life's going...	
ZimHilton	Hilton Barbour	Misty London UK	Networker. Loves travelling - love my 2 daughters more. Perpetually curious.	http://hiltonbarbour.com

Table A-2. An extract of the user profiles contained in the dataset. As it can be seen these profiles comprised the screen name, the user name, location, short biography and website. All of the fields, except for the first two, are optional.

It must be noticed that locations are nothing but free text and, thus, it's up to the user providing a sensible location (e.g. *London* or *NYC*) or something mostly irrelevant (e.g. *at home* or *in the office*). We processed the available locations (62.31% of the users provide such a string) and tried to match them to geographic coordinates by means of a geocoding service[17]. Eventually, 50.36% of the original profiles provided a location string suitable to be matched to actual coordinates.

Given the noisiness of locations, one could argue that some, even many, of the obtained coordinates could be wrong. Nevertheless, as can be seen in the map in Figure A-3 most of the locations must be necessarily correct: users from English speaking countries are majority (the USA and the British Isles are specially prominent); Canada, Australia, New Zealand, Jamaica, Puerto Rico, Netherlands, central Europe and Israel have also a major presence in the sample. Finally, there exist pockets of English-tweeting users in virtually every country but concentrated, as expected, in major global cities (e.g. Paris, Tokyo, Madrid, Seoul, Buenos Aires or São Paulo). On the other hand, the distribution within English speaking countries faithfully corresponds with their population density. So, in short, it seems safe to claim that half of Twitter users provide an accurate geographical location.

[17] http://developer.yahoo.com/maps/rest/V1/geocode.html

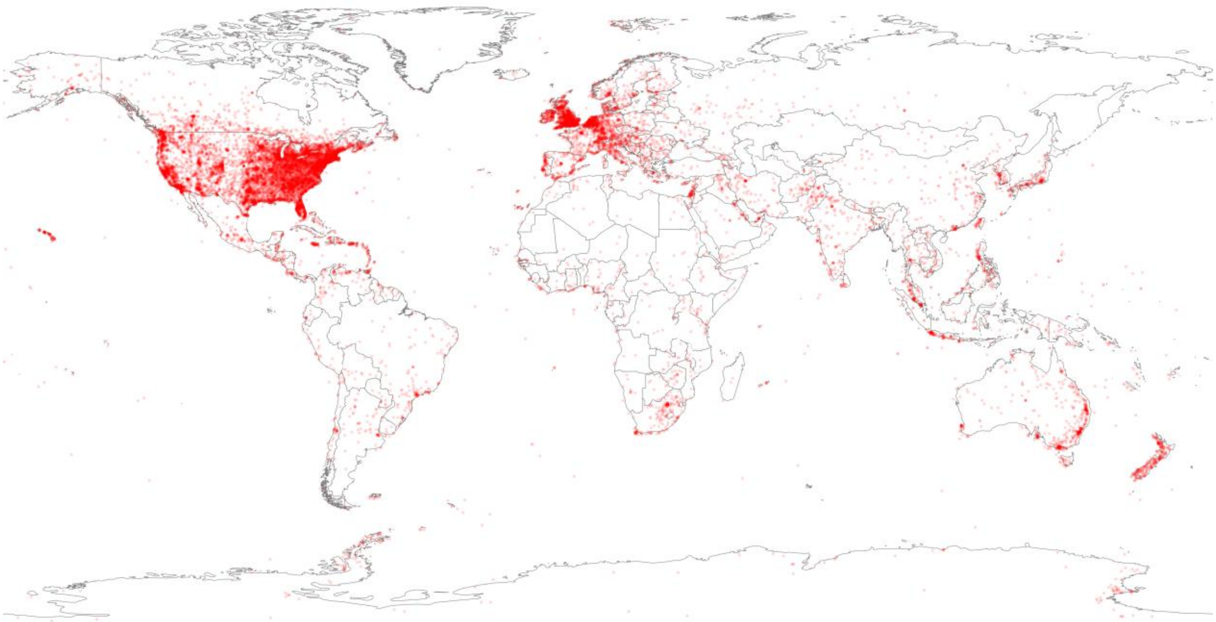

Figure A-3. Map showing the global distribution of users in the dataset. Let's remember that it only contains tweets written in English and, thus, English-speaking countries should concentrate most of the users.

The name and biography fields were in turn employed to infer some demographic features about the sampled users, namely gender and age. To determine the gender we parted from the *"Frequently Occurring First Names and Surnames From the 1990 Census"*[18]; those data files provide 88,799 surnames, 4,275 female first names and 1,219 male first names. We assumed that any user name starting with a first name and ending with a last name from the census was a valid personal name. Certainly, many people provide aliases, just their first name, or their names and/or surnames are not frequent enough to appear in the U.S. Census data; however, we think that this approach is the best for the sake of higher precision.

With regards to those first names appearing in both male and female data files (e.g. *Alexis*, *Charlie*, or *Dominique*) we assign gender according to the frequency of appearance provided the difference was high enough. In this regards, *Alexis* and *Dominique* were always considered female names while *Charlie* was considered a male name. Of course, this is an oversimplification which, certainly, could be improved by taking into account the data in the biography field but we considered that, for the descriptive purposes of this section, it is good enough.

To support that claim some anecdotal evidence can be provided. First, there exists an almost perfect positive correlation between the last name distribution in the U.S. 1990 Census and within the Twitter users (0.9701). The correlation regarding first name lists is smaller but still positive (0.6355 for female names and 0.6356 for male names). Arguably, this can be due to a major presence of young users among *twitterers*. As it can be seen in Table A-3 just one female name appear in both top-10 lists (*Jennifer*) while three are common for male names (*James*, *John* and *Michael*). Both situations seem consistent with the fact that given names have relatively fast turnovers (well under a decade), in particular, female names (cf. [18]).

In addition to this, we computed the most distinctive terms in both male and female biographies by means of a likelihood-ratio test in a way analog to that of [14]. Among the top-10 words for females were *mom*, *girl*, *wife* and *mother*, while *husband*, *guy*, *father*, *dad* and *man* appeared at the top of the list for male users; Table A-4 provides a more exhaustive list.

[18] http://www.census.gov/genealogy/names/

Hence, it seems that our method to assign gender to Twitter users is reasonably accurate and, thus, it can help to provide a picture of the demographics of these users. About 650,000 users provide a personal name (both first and last name appearing the 1990 Census data files), accounting for 36.46% of the users in the graph, from which 58.61% were men and 41.39% women[19].

10 most frequent last names in Twitter	**10 most frequent first female names in Twitter**	**10 most frequent first male names in Twitter**	**10 most frequent last names in the U.S. 1990 Census**	**10 most frequent first female names in the U.S. 1990 Census**	**10 most frequent first male names in the U.S. 1990 Census**
Smith	Sarah	Chris	**Smith**	Mary	**James**
Johnson	**Jennifer**	**Michael**	**Johnson**	Patricia	**John**
Jones	Amanda	**John**	**Williams**	Linda	Robert
Brown	Michelle	**James**	**Jones**	Barbara	**Michael**
Williams	Amy	Mark	**Brown**	Elizabeth	William
Miller	Stephanie	Matt	**Davis**	**Jennifer**	David
Davis	Rachel	**David**	**Miller**	Maria	Richard
Lee	Heather	Mike	**Wilson**	Susan	Charles
Wilson	Katie	Paul	Moore	Margaret	Joseph
Taylor	Jessica	Andrew	**Taylor**	Dorothy	Thomas

Table A-3. List showing the top-10 first and last names in both Twitter and the U.S. 1990 Census. Names appearing in both lists are shown in bold.

Top 50 terms appearing in female biographies			**Top 50 terms appearing in male biographies**		
love	happy	people	**husband**	media	radio
life	friend	art	**guy**	producer	manager
mom	music	beautiful	web	internet	church
girl	world	boys	**father**	marketing	digital
wife	**woman**	single	developer	photographer	**actor**
mother	**mum**	amazing	geek	fan	programmer
friends	married	**wife mother**	**dad**	**husband father**	web developer
lover	**mommy**	**actress**	**man**	pastor	guitar
loves	**sister**	kids	designer	video	ceo
fun	person	books	musician	gamer	enthusiast
live	dance	hi	sports	computer	**dude**
family	heart	love music	technology	business	play
crazy	home	lol	tech	founder	dj
loving	little	dog	director	consultant	online
laugh	want	chick	software	design	beer
fashion	know	teacher	entrepreneur	player	development
daughter	reading		engineer	games	

Table A-4. Top 50 more distinctive terms for female (on the left) and male (on the right) users in Twitter. Those terms which have a clear associated gender are shown in bold. The rest of terms do not have any prior gender but some patterns arise: female users tend to describe their family life, while male users tend to describe their occupations.

In addition to gender, we also tried to determine the age of the users. To that end we relied on the biographies appearing in the profiles, and employed simple patterns (i.e. looking for *"years old"* or *"year-old"* preceded by a number or a numeral). This is also a crude approach, and prone to some mistakes (consider for instance the phrase *"proud mother of 6 years old boy"*) but, all in all, we consider the results highly plausible. Only 10,915 users provide an identifiable "age" in their biographies, the average is 21.67 years with standard deviation of 11.086 years. The distribution is clearly not normal, and users between 15 and 29 years account for 70% of the population. Such a bias toward young adults is, however, expected given the social nature of Twitter, and consistent with the previous discussion about first names.

[19] Although it is out of the scope of this paper, it would be interesting to infer gender for the other two thirds of the users in our dataset using, for instance, the methods described in [22].

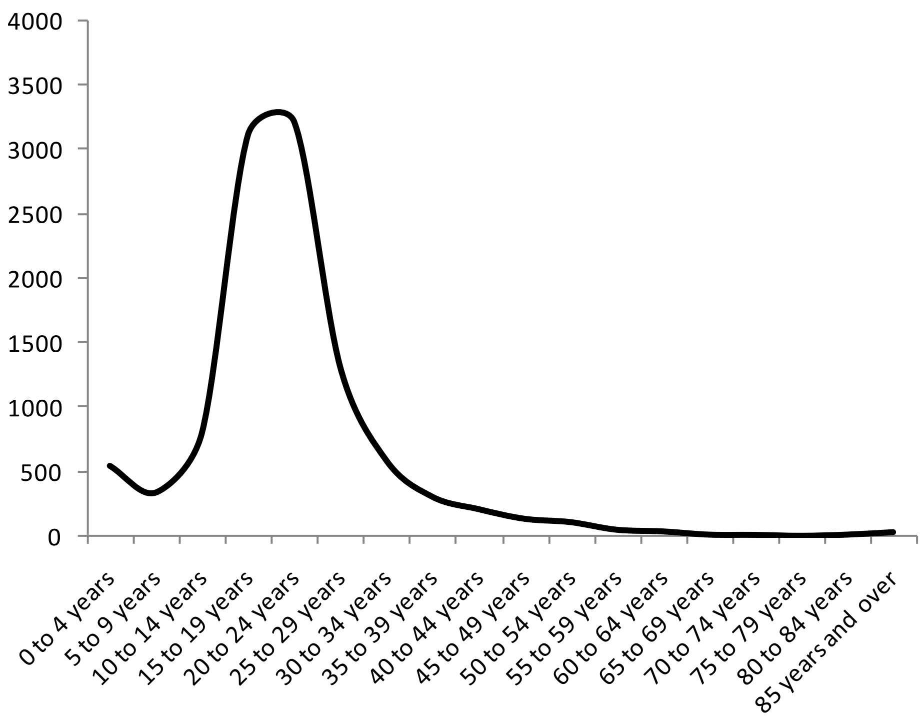

Figure A-4. Age distribution of Twitter users.

As we did with male vs. female users, we also obtained distinctive terms from the users' biographies within each age range. Such terms do provide some clues on the confidence one can place on the aforementioned method to determine the age of Twitter users. Table A-5 shows how the most distinctive terms almost perfectly fit the usual age stereotypes in the U.S. and the UK. For instance, 10 to 14 year-old users are fans of *Twilight*, *Jonas Brothers* and *Miley Cyrus* while users from 15 to 24 years tend to mention *school*, *college* and *university*. Users from 25 to 59 years are commonly married, with kids being especially prominent from 30 onwards. Grandparenthood appears as early as 45-49 years, but is much more common from 55 onwards. Lastly, from 60 to 74 years old, retirement does appear.

Nonetheless to say, there are mistakes: as it was expected, users from 0 to 9 years old are not in fact that age but parents with little children, instead. With regards to the age range of 85 years and more, it seems that most of the users are not indeed that age given the lack of terms appearing in the immediately prior ranges, besides the great heterogeneity of the distinctive terms for that age. If those age ranges are dropped, then the average age is 21.13 years with standard deviation of 9.08 years.

Finally, by combining age and gender it is possible to produce a population pyramid for Twitter (see Figure A-5). Only 4,295 users have both gender and age information and, thus, it's not entirely representative of Twitter users as a whole. However, we can extract some knowledge from it: the current prominence of male users against female users is expected to gradually change towards parity, even, surpassing the number of male users; as it can be seen, female users in the 10-14 and 15-19 ranges (the so-called *digital natives*) clearly surpass the number of male users.

0 to 4 years	**mom**, **son**, **daughter**, **mother**, **boy**, **wife**, home, stay, little, **married**, beautiful, **father**, **home mom**, **twins**
5 to 9 years	**mom**, **son**, **daughter**, **mother**, **twins**, **boy**, **married**, **boys**, **twin**, **wife**, **father**, **dad**, **home**, marketing, **husband**
10 to 14 years	love, girl, name, **twilight**, hey, **jonas brothers**, boy, follow, twitter, **miley cyrus**, grade, **fan**, live, demi, hi
15 to 19 years	music, girl, love, friends, name, **school**, **student**, hey, follow, hi, know, **college**, loves, **jonas brothers**
20 to 24 years	**college student**, life, **university**, **major**, world, **studying**, design, music, working, fun, **graduate**, living, trying
25 to 29 years	**married**, living, work, **wife**, working, male, geek, **mother**, guy, **kids**, writer, teacher, **husband**, manager
30 to 34 years	**married**, **kids**, **mother**, **husband**, **father**, male, lover, single, **children**, beautiful, work, **wife**, female, living
35 to 39 years	**married**, male, **boys**, **children**, single, **mom**, good, **father**, **kids**, **mother**, radio, business, **mum**, guy
40 to 44 years	**married**, single, man, **kids**, **father**, **children**, **mother**, living, woman, **wife**, **mom**, wonderful, male, dogs
45 to 49 years	**grandmother**, **boys**, **wife**, **married**, **kids**, **wonderful boys**, **children**, marketing, **father**, **mother**, **sons**, happily
50 to 54 years	**married**, woman, adult, male, single, **children**, people, **daughters**, man, **father**, cooking, **grandchildren**
55 to 59 years	**married**, cats, **children**, internet, **grandchildren**, hanging, **ex**, man, spiders, surgery, **widow**, male, gardening
60 to 64 years	**retired**, **grandchildren**, south, life, archaeology, flowers, **traveller**, **enjoying**, electro, goals, lawyer, physics
65 to 69 years	man, diving, body, time, **retired**, country, lives, technology, marketing, business, likes, single, school
70 to 74 years	**grandfather**, blogs, running, **retired**, tech, computer, blogger, man, female
75 to 79 years	**granddaughters**, dog, **grandchildren**, stories, bike, grown, tennis, life, living, almost, **children**, play
80 to 84 years	help, **pension**, **retire**, listener, independent, look, money, others, yeah, company, movies, love
≥85 years	newspaper, home, cats, care, aquarium, satisfied, port, bones, river, trades, blogs, books, lol, reader, tea

Table A-5. Distinctive terms from the users' biographies for each age range. Clearly stereotypical terms are shown in bold.

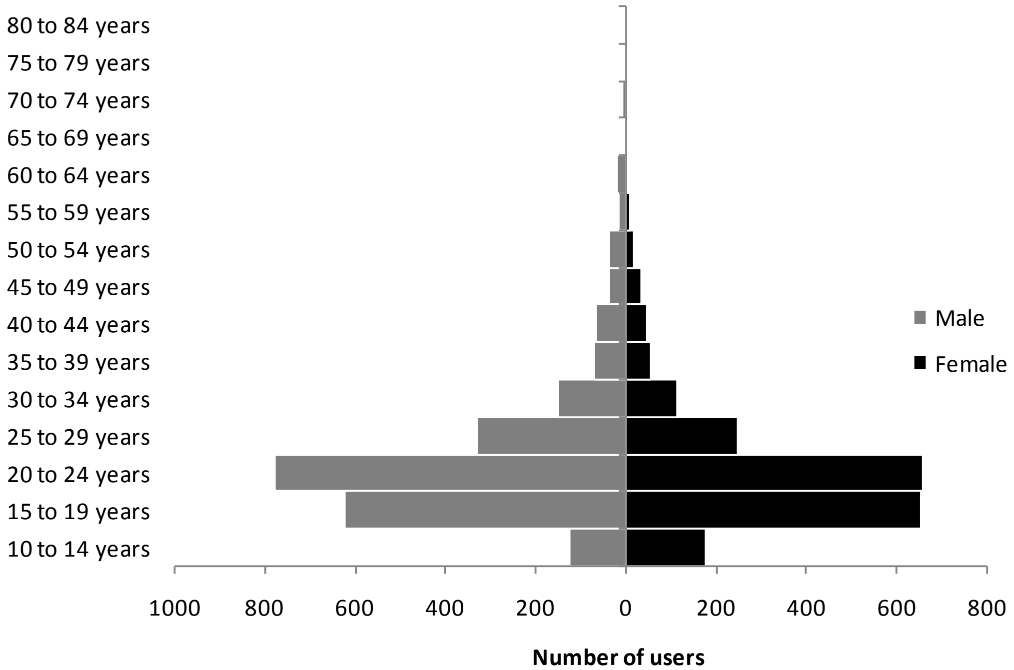

Figure A-5. Population pyramid for Twitter built from the biographies of the 4,295 users who provided both age and gender information.

To sum up, our dataset consists of two different collections: a series of tweets, and the network of users who published them. The tweets collection comprises 27.9 million English tweets posted between January 26, 2009 and August 31, 2009. The user graph consists of 1.8 million users with 134 million connections. 50% of the users are geolocatable, and most of them reside, as expected, in English speaking countries. It was possible to determine gender for one third of the users; of these, 59% were men and 41% women. A relatively small number of users provide age information. Taken them as sample, the average age in Twitter is 21 years and users between 15 and 29 years account for 70% of the total. By examining both age and gender, it seems that the male predominance in Twitter has its days numbered because female surpass male users among the youngsters (10 to 19 years old).