Detecting and Analyzing Automated Activity on Twitter

Chao Michael Zhang[1] and Vern Paxson[1,2*]

[1] University of California, Berkeley, CA
[2] International Computer Science Institute, Berkeley, CA

Abstract. We present a method for determining whether a Twitter account exhibits automated behavior in publishing status updates known as *tweets*. The approach uses only the publicly available timestamp information associated with each tweet. After evaluating its effectiveness, we use it to analyze the Twitter landscape, finding that 16% of active accounts exhibit a high degree of automation. We also find that 11% of accounts that appear to publish exclusively through the browser are in fact automated accounts that spoof the source of the updates.

1 Introduction

Twitter is a microblogging service that allows its members to publish short status updates known as *tweets*. Over 180 M visitors interact with Twitter each month, generating 55 M tweets/day [13]. User accounts and their status updates are public by default, accessible by the general public via Twitter's two application program interfaces (APIs). The large number of users, low privacy expectations, and easy-to-use API have made Twitter a target of abuse, whether relatively benign in the form of spam and disruptive marketing tactics [5], or malicious in the form of links to malware [17] and phishing schemes [8]. Often abuse on Twitter employs automation for actions such as publishing tweets, following another user, and sending links through private messages.

Prior research on Twitter has studied the properties of the social network [10], characteristics of users and their behavior [11], and social interactions between users [9], but not specifically regarding the issue of automation on Twitter (other than our own use of the technique we develop here to assist with finding Twitter "career" spammers [7]). In this work we present a technique for determining whether a Twitter account appears to employ automation to publish tweets, as manifest in fine-grained periodicities in tweet timestamps. Our approaach has the benefit of being able to find legitimate accounts compromised by spammers who employ automation. We evaluate the test's effectiveness and describe its weaknesses, including the ability for determined adversaries to evade it by directly mimicing human posting patterns. Finally, we examine various facets of Twitter as a service and discuss the prevalence of automation in each.

2 Background and Measurement Data

Tweets are short messages (limited to 140 characters) posted to a Twitter account using a browser, a stand-alone application, an API, or SMS messages. Information associated

[*] This work was supported by NSF grants CNS-0831535, CNS-0905631, and NSF-0433702, and ONR MURI Grant N000140911081.

with each tweet includes the time at which the update was created and the source by which the status appears to have been posted. Users on Twitter can subscribe to the tweets of another account by choosing to *follow* that account. The user will then receive that account's tweets through the main "timeline" prominently displayed on the Twitter website and via separate applications, or via SMS messages. Accounts have two main privacy settings: *Public* accounts have their content visible to the general public regardless of whether the visitor is logged in or not, while *protected* accounts can only be viewed by users who have had follow requests accepted by the account owner.

Twitter's "Verified Account" program allows people and companies to show that their account in fact belongs to them. Twitter only makes this program available to a modest number of accounts that deal with mistaken identity or impersonation problems; at the time of this writing there are 1,738 *verified* accounts.

Twitter is a real time communication service, and at any given time there may be certain topics that are widely discussed among members in the community. These *trending topics* are featured prominently to provide users with an up-to-date glimpse at what the community is talking about. Twitter uses algorithms to constantly determine these popular topics, publishes them to the website, and makes them available through APIs.

Twitter provides two APIs through which developers can interact with the service. The "REST API" provides methods for reading and writing data to the main service, while the "Search API" handles queries for searching tweets and obtaining trending topics. The API can be accessed through basic authentication using an account's username and password, or can be accessed through OAuth [2], allowing users to provide third-party applications with access to their data stored on Twitter.

For our purposes we term any account that publishes a significant portion of its tweets automatically using a computer program as a *bot*. We refer to tweets published in real-time by a human as *manual*, or *organic*, tweets.

Data Used in the Study. We draw upon public data associated with accounts and status updates. We evaluated 106,573 distinct accounts using data from 3 weeks in April 2010. Since we rely on public information, we only examine accounts with "public" privacy. For each account, the REST API can return the latest 3,200 tweets, with 200 updates returned per call (we examined a maximum of 300 tweets per account, to avoid skew due to API timeouts). Tweets returned by the API include a timestamp indicating when Twitter received the tweet (1 sec precision), the account's followers and privacy settings, the client program from which the tweet apparently originated, and whether the account has been "verified."

3 Detecting Tweet Automation

We base our detector on the premise that highly automated accounts will exhibit timing patterns that do not manifest in the tweet times of non-automated users. In particular, a human user posting updates to Twitter organically is most likely indifferent towards what second-of-the-minute or what minute-of-the-hour they post updates.[3] Therefore, an organic sequence of update times should appear to be randomly drawn from a uniform distribution across seconds-of-the-minute and minutes-of-the-hour. The upper left

[3] This will certainly be the case if their posting is well-modeled as a Poisson process.

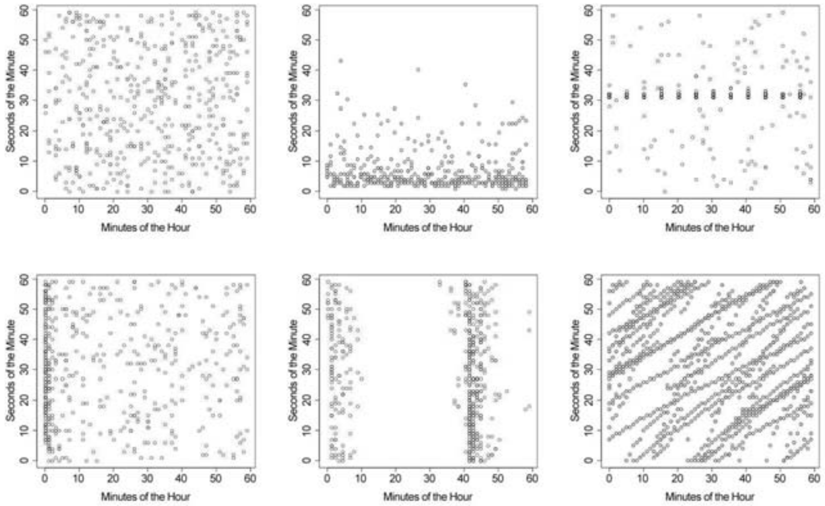

Fig. 1. Timing plots for different Twitter accounts. Each point represents a single tweet. The x-axis gives the tweet's minutes-in-the-hour and the y-axis the seconds-in-the-minute. The upper left plot passes our χ^2 test for expected uniformity, presumably reflecting organic behavior. The others all fail, exhibiting different patterns of non-uniformity, except for the lower right, which exhibits *hyper-uniformity*, too good to be produced by a random-uniform process.

plot in Figure 1 shows a typical *timing graph* for human-generated tweet times. While not completely uniform, they lack noticeable groupings or patterns.

Automated accounts, on the other hand, may exhibit timing distributions that lead to detectable non-uniformity (or excessive uniformity) due to a number of reasons. First, automation is often invoked by job schedulers that execute tasks at specified times or intervals, and these are usually specified in round quantities such as minute-granularity. Furthermore, Twitter imposes a limit of 1,000 tweets/day (as well as finer-grained limits for smaller units of time), so there is no apparent benefit in scheduling automated tweets more often than say one a per-minute basis. Given scheduling at minute-granularity, the seconds-within-the-minute when such tweets appear are unlikely to be uniformly distributed across the minute. The upper middle plot in Figure 1 shows a timing graph of a user who exhibits this type of automated behavior. While the times are distributed somewhat uniformly for minutes-of-the-hour, the user clearly tends to publish updates towards the beginning of the minute.

If scripts publish tweets at scheduled times in each hour, then we will find tweet times clustering at those scheduled minutes. On the other hand, if a script publishes updates on a per-minute basis, it may exhibit a timing pattern that is *too uniform*, which also distinguishes it from organic activity. The upper right plot in Figure 1 shows the timing graph of a user that publishes tweets every 5 minutes in the hour; the lower left plot shows an account that automatically posts updates at the beginning of the hour; and the lower middle plot shows an account that publishes nearly all of its updates during two particular times of the hour.

Non-uniform timing can also arise from delay-based automated behavior: scripts programmed to pause for a certain amount of time after each tweet. Delays that always run the script at the same minutes-of-the-hour will manifest as either extremely non-uniform across minutes-of-the-hour, or, in rare cases, too uniform across minutes-of-the-hour. This latter arises when run times creep into delay-based automation, meaning that small delays that should lead to non-uniformity instead appear to exhibit excessive uniformity. The lower right plot in Figure 1 shows the timing graph of an account that is perfectly uniform across seconds-of-the-minute and minutes-of-the-hour due to what appears to be slowly drifting times. *Thus, we can conclude the presence of automation if we find tweet times either not uniform enough, or too uniform.*

Testing for automated behavior. We use Pearson's χ^2 test to assess whether a set of update times is consistent with the uniform second-of-the-minute and minute-of-the-hour distributions expected from human users. The p-value returned by the χ^2 test is the probability of the observed distribution of times arising if the account is indeed publishing updates uniformly across seconds-of-the-minute or minutes-of-the-hour. If the probability is too low, it indicates that the account exhibits non-uniform behavior in choosing which second-of-the-minute or minute-of-the-hour to publish a post; likewise, if the probability is too high, it suggests that the account is using a mechanism that causes it to publish tweets with a level of uniformity that is unlikely to be observed from natural human use.

For our test we use a two-sided significance level of 0.001, or 0.1%, as the threshold for failing the test. We chose this level after preliminary examination of a small subset of the accounts. We selected a quite low level to avoid incurring many statistical false positives due to the large volume of accounts that we examine. Thus, we expect only 2 in 1,000 human accounts with uniform distributions to fail each test.

A common rule of thumb for Pearson's χ^2 test is that 80% of bins should have an expected count of at least 5 [6]. Therefore if we have 300 timestamps for an account we use 60 bins for assessing seconds-of-the-minute and minutes-of-the-hour. If we have fewer, then we use only 6 bins, unless the account has fewer than 30 tweets, in which case we exclude it due to insufficient data. Eliminating such accounts does not significantly impair our study as we presume that the interesting uses of automation occur when accounts regularly tweet.

Automated accounts can exhibit non-uniform timing patterns for both seconds-of-the-minute and minutes-of-the-hour, both indicative of automation. Therefore, we perform a separate χ^2 test for each, with a failure of either indicating automation.

4 Evaluating the Test

An important issue is that we lack ground truth regarding whether accounts are truly automated or organic, and also whether automation reflects unwanted activity. However, we form a partial assessment as follows. From an initial evaluation of 18,147 accounts we found that 975 accounts had seconds-of-the-minute p-values less than 0.001, and 15 accounts had p-values greater than 0.999. The same figures for minutes-of-the-hour are 2,599 p-values less than 0.001 and 76 greater than .999.

We manually examined hundreds of timing graphs to confirm they exhibited clear non-uniform or hyper-uniform behavior, and randomly selected dozens of accounts for manual verification. (Accounts that did not visibly manifest non-uniform behavior, but were flagged by the test, generally turned out to indeed use third party applications that automate tweets.) This latter included an examination of the user's profile and their first page of recent status updates. In nearly all cases we could determine that the account exhibited strong evidence of likely automation not reflecting social human use, based on status updates (i.e., number of updates, sources, frequency, and contents) and other features of the account's Twitter page (i.e., user icon, background image, screenname, number of followers and friends, and website URL). See below for further discussion of our evaluation of false positives and false negatives.

This assessment gives us confidence that a significance level of 0.001 can effectively capture accounts that exhibit anomalous timing behavior. However, we also note that such a stringent significance level can cost us the opportunity of observing *hybrid* accounts that publish with a mix of manual and automatic updates. Some hybrid users may utilize different applications for these two kinds of updates, allowing us to separate these sources in order to evaluate our test. For example, one hybrid we identified used the third-party applications TweetDeck [3] and HootSuite [12], both applications that provide an interface for reading and creating tweets. However, TweetDeck does not offer functionality for automating tweet creation, while HootSuite provides a scheduling feature. This account's timing graph exhibits distinct periodicity. Testing only the tweets posted from TweetDeck, however, does not exhibit such patterns (and passes the χ^2 test), while tweets originating from "HootSuite" exhibit updates at five minute intervals, failing the χ^2 test.

False Positives. A false positive occurs an account fails our test but is in fact organic. Along with statistical fluctuations (which will contribute about 2 false positives per 1,000 accounts we assess), these can arise due to legitimate organic use that deviates from uniform timing. For example, a student who only publishes Twitter updates in between class periods may fail our test because their tweets will tend towards certain minutes-of-the-hour.

An example of an account that fails our test but otherwise appears to be organic is the account of television personality Phil McGraw, also known as Dr. Phil [1]. After inspecting the account, we found that it consistently publishes one update per day shortly before the show begins to remind followers to watch. Although these updates are manually generated, they are skewed towards the first half of the hour.

While we discovered a few false positives along these lines, we note that all of them concerned accounts that failed on minutes-of-the-hour for the type of reason described above. We have not discovered any apparently legitimate human account that exhibits anomalous timings for seconds-of-the-minute.

False Negatives. On the other hand, our false negative rate is likely considerably higher for a number of reasons. First, as discussed above, hybrid behavior can mask automated posting due to blending it with organic posting. We could potentially detect more such instances by using a less stringent significance level, but at the cost of more statistical false positives. Second, automated accounts that exhibit uniformity in some fashion will of course be missed by our test. In particular, one form of this can arise from *copycat automation*, i.e., an automated account that posts in reflection of non-automated timings. For example, an automated accounts triggered by an RSS feed will reflect the timings of the source rather than a specific schedule.

Evasion. One can easily design an automated account to evade the χ^2 test by uniformly spreading its tweets across seconds-of-the-minute and minutes-of-the-hour. For example, the account could post whenever a known-organic account posts; or simply generate exponentially distributed interarrivals. There does not seem to currently exist any incentive for automated accounts to be intentional about exhibiting uniformity. However, if Twitter adopts a test like ours as a countermeasure to detect possible abuse, then accounts may begin evading the test in this way.

5 Analyzing Twitter's Landscape

Using the χ^2 test, we analyzed public tweets and accounts to determine the prevalence of automated accounts on the service and how the use of automation varies with respect to different factors. We sampled the public timeline of global tweets via the REST API, which makes available the 20 most recent tweets, refreshed every minute. We were therefore able to obtain a sample of 1,200 tweets per hour. In addition, we used the Search API to query for samples based on keywords and to obtain trending topics. For a range of keywords, we performed a search every minute and recorded the accounts behind the 10 most recent results, for which we then analyzed the posting account. We sampled search results for between two and four days for each keyword. In addition to the constantly changing public timeline and sampled search results, we also obtained accounts from various static lists, including verified users, most-followed users, and followers of the most-popular account, collecting up to 300 tweets for each account.

For each account we have six possible dispositions. *Passed* accounts pass the χ^2 test while *Failed* accounts do not. *Insufficient* accounts do not have the 30 status updates necessary to perform the test. *Protected* accounts have their privacy settings set to protected, so we could not test them. *Suspended* accounts have been suspended by Twitter for reasons such as spamming and abusing the API. These accounts are rendered completely inaccessible through the API. However, their user IDs may persist for a time in various places on Twitter, and therefore may be included in our analysis. *Not Found* accounts no longer exist on Twitter. When an individual or business deactivates their Twitter account, the API returns an error when requesting data from that account. However, the user ID may persist on various pages of Twitter for up to 30 days, and may be detected by our analysis.

Table 1 summarizes our results. We note that accounts might exhibit varying degrees of automation depending on temporal factors such as time of the day or day of the week. For example, an account may syndicate news from a news source that publishes more heavily during the waking hours of the day, or may publish from a source that is inactive on weekends. Therefore, a more accurate assessment of automated activity on Twitter may monitor activity over the course of weeks or months in order to determine

Table 1. Automation testing results for different facets of the Twitter landscape (lower bounds)

Facet	Total	Passed	Failed	Insufficient	Protected	Suspended	Not Found
Public timeline accounts	19,436	15,330	2,817	1,176	66	47	0
Public timeline tweets	18,331	14,790	2,475	983	59	24	0
Verified users	1,738	1,531	113	66	17	6	5
Most followed (all)	1,000	862	121	15	1	0	1
(verified)	400	373	25	2	0	0	0
(not verified)	600	489	96	13	1	0	1
Trending topics	14,230	13,260	617	286	58	8	1

average levels of automation. Our present analysis does not take these considerations into account, which we leave for future work. Finally, we emphasize that our estimates likely reflect lower bounds, as we will overlook both low-rate automation (too few samples to apply the χ^2 test) and automation that already employs randomization to avoid appearing regular.

Public Timeline. The Twitter public timeline provides a sample of the thousands of tweets being sent via the service each minute. Thus, we can use it to estimate the prevalence of automation for public statuses on Twitter overall. The *Public timeline accounts* line of Table 1 reflects a sample from two days in April 2010. Of the 19,436 accounts examined during this period, we could test 18,147 using our χ^2 method. We find that 16% of the accounts publishing tweets exhibit discernible automation.

A study conducted in August 2009 analyzed 11.5 million accounts, classifying those publishing $>$150 updates per day as bots [15]. The report concluded that at least 24% of all tweets were generated by automated bots. Around this time, Twitter began to focus on reducing spam in the service, and in March 2010 published the claim that the tweet spam rate had fallen below 1% [5]. To test these claims, we also ran a separate analysis (on different, somewhat smaller data) of the public timeline weighted by tweet rather than by account (*Public timeline tweets* row). We find that 14% of public tweets come from automated sources, suggesting that Twitter has indeed reduced the amount of unwanted automation on the service (if the methodology used by [15] has an accuracy comparable to ours). However, unless the vast majority of these automated tweets are not spam, our results also indicate that the problem of spam is still far from being solved.

Verified Users. That verified accounts are often owned by celebrities and popular companies (and Twitter manually approves accounts in the program) argues against these accounts exhibiting strong automation in their tweets. A heavily automated account may reflect badly on fans and customers, and would likely be harder to have approved by Twitter. The *Verified users* row in Table 1 shows the results of our analysis of these accounts. We find that 6.9% failed our test—the amount of automation seen in verified accounts is indeed less than the proportion in the general Twitter population. Among the verified accounts that failed were: (1) popular bands reminding fans of concerts and TV appearances, (2) TV shows reminding their fans of episodes each day, (3) political figures and parties publishing links to news articles, (4) journalists publishing links to their organizations, (5) non-profit organizations sharing links to issues around the world, and (6) government organizations publishing news and alerts to

Table 2. Profiles of different sources used to publish tweets.

	Overall	Automation	Bot	Bot	Organic	Organic
Source	Use	Rate	Rate	Exclusivity	Rate	Exclusivity
Web	31%	6.4%	11.8%	85%	37%	82%
Ubertwitter	9.4%	2.3%			11.9%	87%
Twitterfeed	7.5%	62.0%	27.8%	94%	3.7%	95%
Tweetdeck	6.6%	3.9%	1.5%	76%	8.2%	77%
REST API	5.9%	60.0%	21.0%	96%	3%	92%
Echofon	4%	2.1%			5%	77%
Mobile	2%	1.9%			2.5%	73%
Tweetie	1.6%	3.0%			2%	73%
Txt	1.6%	2.6%			2%	75%
Hootsuite	1.4%	51.0%	4.1%	84%		

the public. Thus, common reasons for verified accounts failing our test were that they syndicated news, shared links, or sent reminders to followers in an automated way.

Most Followed Users. Although Twitter does not publish a list of most-followed users, certain 3rd-party websites do. Using the list provided by TwitterCounter [16], we analyzed the 1,000 most-followed accounts on Twitter. We find that 12% of the testable accounts failed our χ^2 test (*Most followed (all)* row). Only 6.3% of the verified accounts (next row) failed, slightly lower than the 6.9% found when analyzing all verified Twitter accounts. Of the remaining 600 not-verified accounts, significantly more (16%) were likely to be automated. Manually examining the 96 non-verified accounts that failed, many of them were news websites, blogs, and TV shows that use Twitter to broadcast new content to followers.

Trending Topics. Twitter publishes a constantly updated list of the 10 most popular words or phrases at any given time, providing users with a realtime glimpse at the topics being discussed by the Twitter community. Since many users follow trending topics by reading the latest tweets that contain those particular terms, it would seem profitable for automated accounts to target currently trending topic keywords. To test the trending topics for automation, we performed a search for the first trending topic once per minute, and tested the accounts behind the resulting tweets. As the results Table 1 show, we found that only 4.7% of accounts participating in the trending topic discussions on Twitter exhibited strongly automated behavior—significantly less than the 16% automation found in the public timeline.

This lower rate of automation may indicate that Twitter is careful in preventing automated tweets from polluting the trending topic discussions, since the tweets posted in response to trending topics are frequently viewed by both members and visitors. Alternatively, perhaps the number of human users is simply proportionally higher in searches for trending topics compared to the public timeline, or spammers have not widely adopted this tactic yet.

Keyword Search Results. Using the Twitter Search API, we evaluated the accounts behind the search results for 24 keywords that we believed might result in varying levels of automation. (Our aim here is to obtain a qualitative sense of automated-vs.-non-automated topics, rather than a representative assessment.) Sorted in descending order by the proportion of testable accounts that appear automated, the words were: *mortgage* **(48%)**, *jobs*, *insurance*, *news*, *discount*, *free*, *money* **(31%)**, *click*, *sex*, *poker*, *photography* **(24%)**, *video*, *download*, *bot*, *video*, *viagra* **(17.5%)**, *porn*, *school*, *tv*, *bieber*, *jesus* **(8.3%)**, *happy*, *bored*, *god* **(5.0%)**.

Most keywords tested had automation rates higher than the global 16% automation rate, particularly keywords commonly associated with spam ("discount", "free", "sex", "poker", and "download"). Likewise, keywords with lower automation rates often reflect terms not commonly associated with spam ("jesus", "happy", "bored", "god"). It is surprising though to find that "photography" had a higher rate of automation than "viagra". However, manually searching these keywords indeed reveals a significant amount of automated linking to photography-related articles and websites, while "viagra" often appears in lighthearted messages or jokes posted organically.

A more comprehensive study might directly analyze the frequencies of words that appear in the updates of automated/organic accounts. We leave this for future work.

Tweet Sources. For each account tested, we also analyzed the source appearing most often in that account's tweets. Table 2 summarizes the usage of the most popular sources. *Overall Use* is the percentage of tweets we examined that used the given source. *Automation Rate* is the proportion of those tweets belonging to accounts that we identified as automated. The next two columns reflect what proportion of automated accounts used the given source, and of those, how many used *only* that source ("Exclusivity"). The final two columns summarize the same information for non-automated accounts. Empty table entries reflect that the given entry corresponded to marginal activity (not in the top ten sources for either bots or organic activity, respectively).

We see sharp differences in usage patterns depending on the sources employed. Activity from Twitterfeed, REST API, and Hootsuite is very often automated, while other sources exhibited automation rates far below the overall average rate of 16%. Indeed, many of the services favored by organic users (e.g., UberTwitter [4], TweetDeck [3], and Echofon [14]) do not offer any scheduling features. This suggests that consideration of publishing source might prove beneficial for identifying unwanted/malicious Twitter activity. However, just about all of the top sources are also used organically, so we cannot simply filter by source without considering other factors.

Based on these findings, a possible way to improve our testing would be to examine the publishing times of each of an account's sources separately. Doing so might readily identify both hybrid accounts and *hijacked* accounts for which an attacker usurps use of what is otherwise a legitimate, organic account.

6 Summary

We have presented a method for detecting instances of automated Twitter accounts using only the publicly available timestamp associated with each of an account's tweets. We find that automated accounts exhibit distinct timing patterns that we can not only observe visually, but also detect in a mechanized fashion using Pearson's χ^2 test.

Testing 19,436 accounts from the public timeline, we find 16% exhibit highly automated behavior, and that 12% of automated accounts spoof their tweet source as "web," apparently to appear organic. (Note that these at best reflect evasive postings, because legitimate automation would presumably use the API rather than a web browser.) We also find that verified accounts, most-followed accounts, and followers of the most-followed account all have lower automation rates than the public timeline (6.9%, 12%, and 4.2%, respectively). Trending topic search results were found to have a lower rate as well, with 4.7% automation. We also find that keywords more associated with spam generally have higher automation rates than other keywords. We also examined the apparent source of tweets, finding that automated sources utilize services that provide automation and scheduling, while organic users often use Twitter's web interface or other non-automated services.

A practical application of our methodology could be to use it in conjunction with existing spam prevention measures such as community flagging of inappropriate or abusive accounts. The ability to quickly assess that an account operates in an automated fashion would allow operators to expedite paying attention to such complaints, allowing them to more quickly and effectively combat cases of serious spam and other abuse.

References

1. Dr. Phil (DrPhil) on Twitter. `http://twitter.com/DrPhil`, 2010.
2. OAuth FAQ. `http://apiwiki.twitter.com/OAuth−FAQ`, 2010.
3. TweetDeck. `http://www.tweetdeck.com/`, 2010.
4. UberTwitter. `http://www.ubertwitter.com/`, 2010.
5. A. Chowdhury. State of Twitter Spam. `http://blog.twitter.com/2010/03/state−of−twitter−spam.html`, 2010.
6. R. B. D'Agostino and M. A. Stephens, editors. *Goodness-of-Fit Techniques*. Marcel Dekker Inc., 1986.
7. C. Grier, K. Thomas, V. Paxson, and M. Zhang. @spam: The Underground on 140 Characters or Less. In *Proc. ACM CCS*, 2010.
8. D. Harvey. Trust And Safety. `http://blog.twitter.com/2010/03/trust−and−safety.html`, 2009.
9. B. A. Huberman, D. M. Romero, and F. Wu. Social networks that matter: Twitter under the microscope. Technical report, Social Coputing Laboratory, HP Labs, 2008.
10. A. Java, X. Song, T. Finin, and B. Tseng. Why We Twitter: Understanding Microblogging Usage and Communities. In *Proc. Joint 9th WEBKDD and 1st SNA-KDD Workshop*, 2007.
11. B. Krishnamurthy, P. Gill, and M. Arlitt. A few chirps about Twitter. In *Proc. ACM SIGCOMM Workshop on Online Social Networks*, 2008.
12. I. Media. HootSuite. `http://hootsuite.com/`, 2010.
13. C. Miller. Twitter Makes Itself More Useful. `http://bits.blogs.nytimes.com/2010/04/14/twitter−makes−itself−more−useful/`, 2010.
14. naan studio. Echofon. `http://www.echofon.com/`, 2010.
15. Sysomos. Inside Twitter: An In-Depth Look at the 5% of Most Active Users. `http://sysomos.com/insidetwitter/mostactiveusers`, 2009.
16. TwitterCounter.com. The 1000 most popular Twitter users. `http://twittercounter.com/pages/100/`, 2010.
17. K. Zetter. Trick or Tweet? Malware Abundant in Twitter URLs. `http://www.wired.com/threatlevel/2009/10/twitter_malware/`, 2009.