First Monday, Volume 18, Number 7 - 1 July 2013

PEER-REVIEWED JOURNAL ON THE INTERNET

You followed my bot! Transforming robots into influential users in Twitter

by Johnnatan Messias, Lucas Schmidt, Ricardo Oliveira, and Fabrício Benevenuto

Abstract

Systems like Klout and Twitalyzer were developed as an attempt to measure the influence of users within social networks. Although the algorithms used by these systems are not public known, they have been widely used to rank users according to their influence, especially in the Twitter social network. As media companies might base their viral marketing campaigns on influence scores, users might attempt to boost their influence scores with simple mechanisms like following unknown users to be followed back or even interacting with those who reciprocate these actions. In this paper, we investigate if widely used influence scores are vulnerable and easy to manipulate. Our approach consists of developing Twitter bot accounts able to interact with real users to verify strategies that can increase their influence scores according to different systems. Our results show that it is possible to become influential using very simple strategies, suggesting that these systems should review their influence score algorithms to avoid accounting with automatic activity.

Contents

1. Introduction

Social networks have become quite relevant and increasingly important in modern society. Among them, Twitter has obtained a great number of users. Recent estimates suggest that over 200 million active Twitter users post 150 million tweets (messages) daily (Kirkpatrick, 2011). If Twitter was a country, its active users would rank it as the world's fifth largest.

In systems like Twitter, users can influence and be influenced by others, and that fact has attracted political interest and attention from marketing corporations. In this context, several companies became specialized in measuring influence in Twitter and other social networks. Among the most popular systems, we cite Klout (http://klout.com/) and Twitalyzer (http://twitalyzer.com). Both use simple methods to measure influence, without revealing the details of their algorithms to the public.

This paper addresses these questions via the analysis of the mobile photo–sharing application Instagram, a social network that offers its users a way to upload photos, apply different manipulation tools ('filters') to transform the appearance of an image, and share them instantly with the user's friends (using Instagram's app or other social networking sites such as Facebook, Foursquare, Twitter, etc.) [1]. First launched in October 2010, as of February 2013 the application already had over 100 million registered users who have shared nearly four billion photos from all over the globe [2].

Several newspapers and other sources (Purohit, *et al.*, 2011; Anger and Kittl, 2011; Brown and Feng, 2011; Yan and Kaziunas, 2012) have been using such tools as attempts to produce rankings of influential users. As an example, an article in the *New York Times* presented a study about the most influential people in the world, based on Twitalyzer (Leonhardt, 2011). Among the most influential ones, the study included Brazilian comedian Rafinha Bastos, American rapper Snoop Dogg, and the President of the United States, Barack Obama.

The popularity of these social influence metrics raises some concerns about how these approaches work and how vulnerable they are to attacks. As shown in recent research (Ghosh, *et al.*, 2012b), several users in Twitter seek to amass social capital and influence in the network in order to leverage it to promote their tweets. So it is natural that they would interconnect with others having a similar desire to amass social capital. Thus, these users might collude with each other, following and being following, retweeting and being retweeted, making some influence scores vulnerable.

In this paper we investigate if simple automated strategies can turn common users into influential ones, according to the rankings of Klout and Twitalyzer. To do that, we created simple robots able to interact through Twitter accounts, as if they were real Twitter users, by exchanging information, as well as following and gaining new followers over a period of 90 days. The result of the analyses has shown that the Klout and Twitalyzer tools are vulnerable to simple automatic strategies, as our robots were able to gain considerable influence scores. It is important to highlight that even spammers could pretend that they are celebrities with a high level of influence based on Klout Score and Twitalyzer Impact.

The rest of this paper is organized as follows. In section 2, we present a summary of research related to the subject of this article. Section 3 presents robot algorithms employed to interact in the social network. In section 4, we show experimental scenarios. In section 5, we present a method for collecting data as well as the bot execution log. In section 6, we present the results. Finally, in section 7 we present conclusions and directions for future work.

2. Related work

The theory of influentials assumes that a minority of members in a society possess qualities that make them exceptionally persuasive in spreading ideas to others. Thus, by identifying and convincing a small number of influential individuals, a viral campaign can reach a wide audience at a small cost. This theory has spread well beyond academia and has been adopted in many marketing strategies (Gladwell, 2000; Keller and Berry, 2003; Katz and Lazarsfeld, 2006; Rogers, 2003). We argue that many Twitter users might make use of artificial strategies to boost their influence scores to spread their own information or even to become the target of the marketing industry.

Measuring influence in Twitter is a lengthy and complex task. Lengthy because it is an analysis that needs to account past actions of users, and complex because there is no precise consensus as to what influence is and how to measure it. Some studies indicate that a person's influence in Twitter is more related to the propagation and repercussion of tweets than the number of followers that a user has (Cha, *et al.*, 2010). There have been several other attempts to create methods for properly measuring a user's influence on Twitter (Bakshy, *et al.*, 2011; Lee, *et al.*, 2010; Romero, *et al.*, 2011). Weng, *et al.* (2010) proposed TwitterRank, a method which uses both the Twitter connections graph and information of published tweets, in order to identify influential users. Pal and Counts (2011) used grouping and classification of more than 15 characteristics extracted from Twitter graph and from tweets posted by users, for identifying the most influential ones. More recently, Ghosh, *et al.* (2012a) proposed a system namely Cognos that allows one to search for experts in topics in Twitter. The idea behind the tool is based on crowdsourcing information extracted from Twitter Lists, where users annotate others with descriptions of expertise. Authors validate Cognos by comparing it with Twitter whom–to–follow system in a blind test, where users choose which system is better.

Overall, although these methods consider more elaborated strategies, it is unclear if they are deployed in current influence score systems like Klout and Twitalyzer. Klout consists of a score based on 25 measures divided into three groups in order to arrive at the Klout Score: Network Impact, Amplification Probability and True Reach. In the case of Twitalyzer, it computes a score named Twitalyzer Impact based on 15 measures similar to the used by Klout. Details of how exactly these scores are computed are not unveiled by the companies.

Ghosh, *et al.* (2012b) investigated for the first time the activity of link farms in Twitter (*i.e.*, groups of users trying linking to each other to increase to amass social capital). They found that a small group of legitimate users who are both popular and active on Twitter, are accountable for the majority of link farms. They seek to gather social capital by following several users and gaining new followers. Spammers take advantage of such method in order to acquire followers and reputation on Twitter. As a way to reduce the influence of this sort of behavior, Ghosh, *et al.* (2012b) proposed a ranking scheme, known as Collusionrank, where users are penalized for following spammers, reducing the influence of spammers and their followers.

Overall, these efforts are complementary to our work. Our basic research question is simply: *Can simple strategies turn routine Twitter users into influentials, according to popular influence score systems?*

3. Bots construction

Before explaining how the bots work, let us introduce them. Two bots were created by using the Twitter API Python. The first bot, named *fepessoinha* (https://twitter.com/fepessoinhas2) just follows and acquires other users automatically. In order to follow users our algorithm initially follows a random user and recursively follows 30 other random users extracted from the list of followees of the collected user. This process is repeated, until the bot reaches the limit of 2,000 followed users, which is a limit imposed by Twitter to ordinary users. Figure 1 exhibits a profile photo of the fepessoinha bot [1].

Figure 1: *Fepessoinha* bot profile — Twitter.

The second bot, named *scarina* (https://twitter.com/scarina91), is able to interact as if it was a common user, by posting tweets in addition to following other users. To follow users, this bot uses a similar mechanism implemented in the bot *fepessoinha*, with two modifications. First, after following those 2,000 users, the algorithm unfollows the users who did not follow the bot back after one day. Second, *scarina* posts tweets regularly about a specific theme, containing popular topics. To do this, the bot reads a dictionary of words related to the most popular open TV channel in Brazil, Rede *Globo* [2]. We inserted in this dictionary words extracted from the most accessed and searched terms on the TV channel Web site. These terms consist of names of Brazilian celebrities, soap operas, or simple keywords like "news" and "highlights". In total we used 30 terms.

The *scarina* algorithm search in Twitter sentences like "Globo + **word**", in which the word is randomly selected from this dictionary. Afterwards, *scarina* selects the four most recently posted tweets from the search result to retweet or to repost as if it was an original tweet. The time between posts randomly varies from 0 seconds to one hour to avoid posts in bursts. This entire process is then repeated over and over. Figure 2 exhibits the *scarina* bot profile photo.

Figure 2: *Scarina* bot profile — Twitter.

Both bots constantly checked the following Twitter API restrictions to guarantee they would not be identified as bots and blocked by Twitter:

- Maximum number of requests per hour is 350;
- Each user can follow up to 2,000 users;
- A user cannot follow more than 1,000 users per day.

4. Experimental scenarios with bots

Four experimental scenarios were created to analyze how the bots acquire influence according to Klout and Twitalyzer. These scenarios test the vulnerability of influence classification methods and how easily they can be manipulated.

- Scenario 1: Following users (*fepesoinha* bot): The *fepessoinha* bot followed the maximum limit of users allowed by Twitter (2,000) via API, trying to gain followers.
- Scenario 2: Following and keeping only those who followed it back (*scarina* bot): In this scenario *scarina* follows the maximum number of users allowed by Twitter and unfollows those who did not followed it back.
- Scenario 3: Posting tweets (*scarina* bot): In this scenario we evaluate what happens with influence scores as *scarina* begins to post tweets automatically.
- Scenario 4: Checking interruptions is posts (*scarina* bot): Lastly we evaluate what happens with the bot influence when *scarina* stop posting tweets and stays inactive for a certain period of time.

5. Bots execution and monitoring

In order to evaluate the aforementioned scenarios, we executed our bots from 2 September 2011 to 2 December 2011 using two machines located at Universidade Federal de Ouro Preto (UFOP), Brazil. To monitor and record new followers acquired by these accounts as well as mentions and retweets, we set up their Twitter accounts so that every single action (*e.g.*, a retweet) would be reported by e–mail to an accounted created for this specific purpose. With this strategy, we were able to record all actions and activities in Twitter that involved the bots during the 90 days of the experiment. Subsequently, we created a parser algorithm to collect the e–mail messages in ".eml" file downloaded from the Yahoo! e–mail service. More specifically, we were able to obtain the following Twitter information: favorite tweets, quoted tweets, messages, answered tweets, retweets and follower users.

After the execution process, we assessed the Klout and Twitalyzer sites in order to collect results of all interactions. At the Klout site it was possible to collect influence scores for each day of experiment, by means of graphics. From the Twitalyzer site, we were able to collect only the final influence score for each bot.

6. Results

We evaluated the Klout score and the Twitalyzer Impact as a function of each experimental scenario. Both scores vary from 0 to 100.

6.1. Scenario 1

Figure 3 shows in detail the number of followers that the bot *fepessoinha* acquired, resulting in 417 followers. The number of followers was highest in the period in which the bot *fepessoinha* executed the process to follow 2,000 users (during the first nine days).

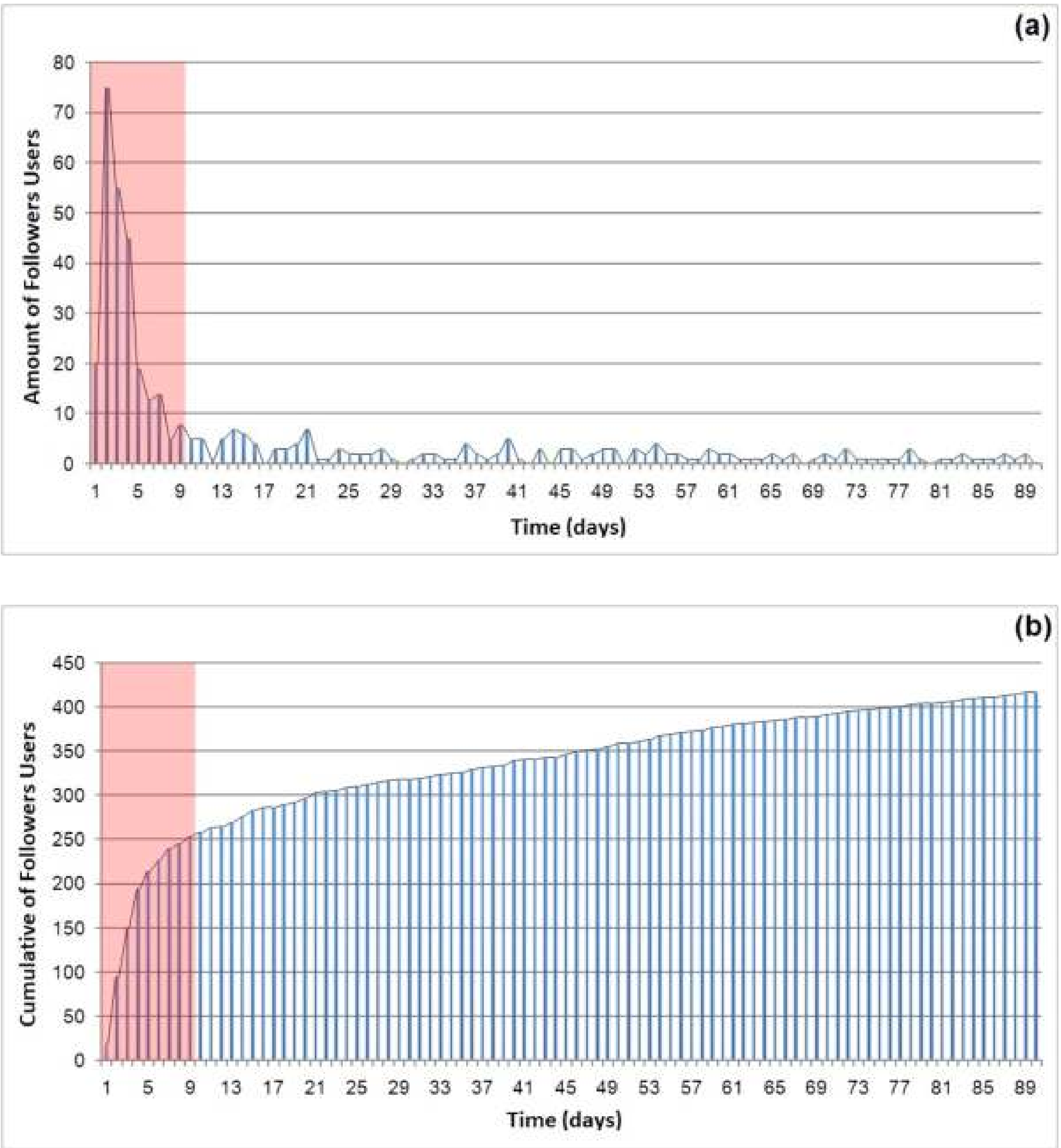


Figure 3: (a) Followers acquired by *fepessoinha*; (b) Cumulative number of followers of *fepessoinha*.

In Figure 4, we contrast the Klout Score of *fepessoinha* for the same period. The Klout Score shows the final result of one user's influence. It is possible to see that in the first two days there was a rapid increase due to new users who followed the bot, reaching a value of 18 in the Klout Score. In the end of the execution of the bot's algorithm (after nine days), we note that the Klout Score diminishes until it reaches a value of 12.3 at the end of the 90–day experiment. Similarly, the bot reached a value of nine in Twitalyzer, for the 90 days of the experiment.

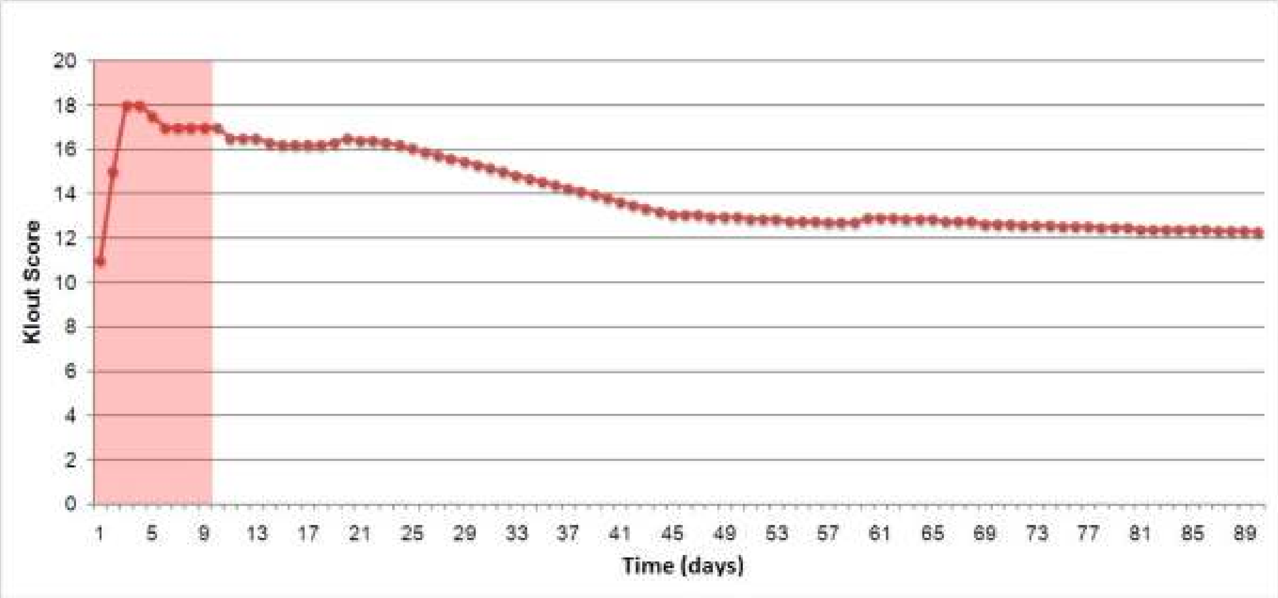
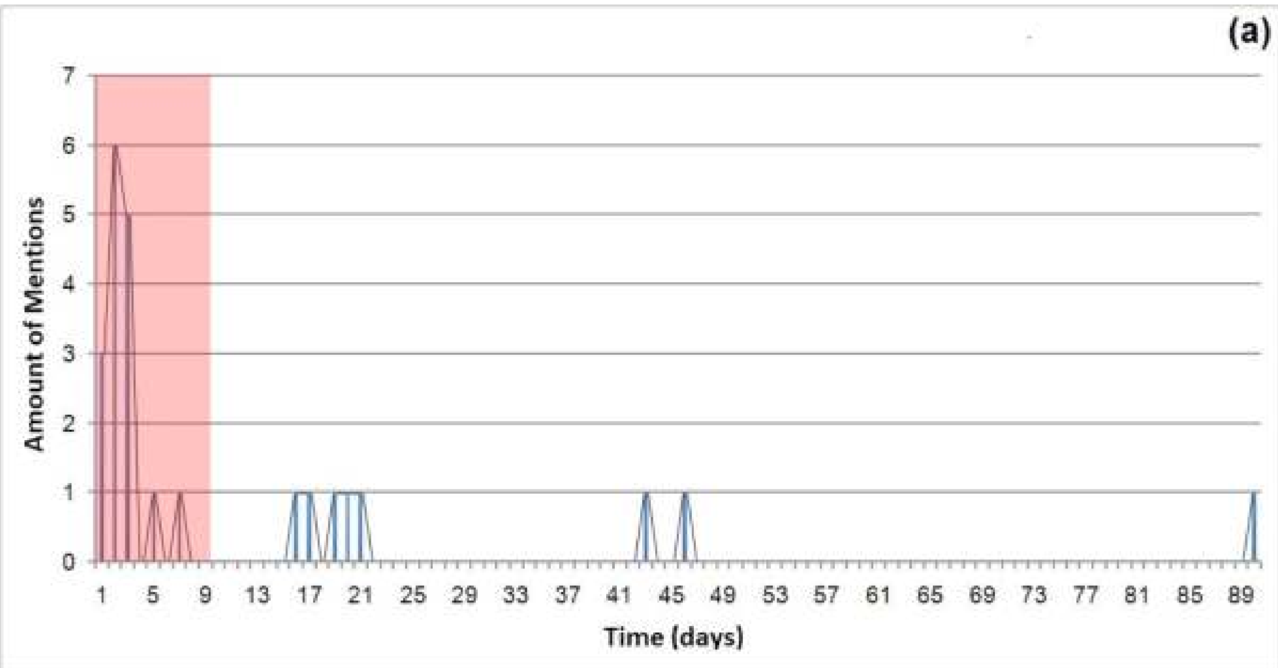


Figure 4: Bot *fepessoinha's* Klout Score result.

During the 90 days of the experiment, *fepessoinha* obtained 0 favorite tweets and 0 retweets, due to the fact that it published only seven tweets (posted manually just after the bot account was created). According to Figure 5, the bot had a considerable number of mentioned tweets (24 in total) and private messages received (21 in total) by the end of the experiment. We observed that a large fraction of the mentions and messages occurred in the first nine days, when the bot algorithm was executed. Most of the mentions and messages were questions aiming at understanding who *fepessoinha* was, and from where the users may know it.



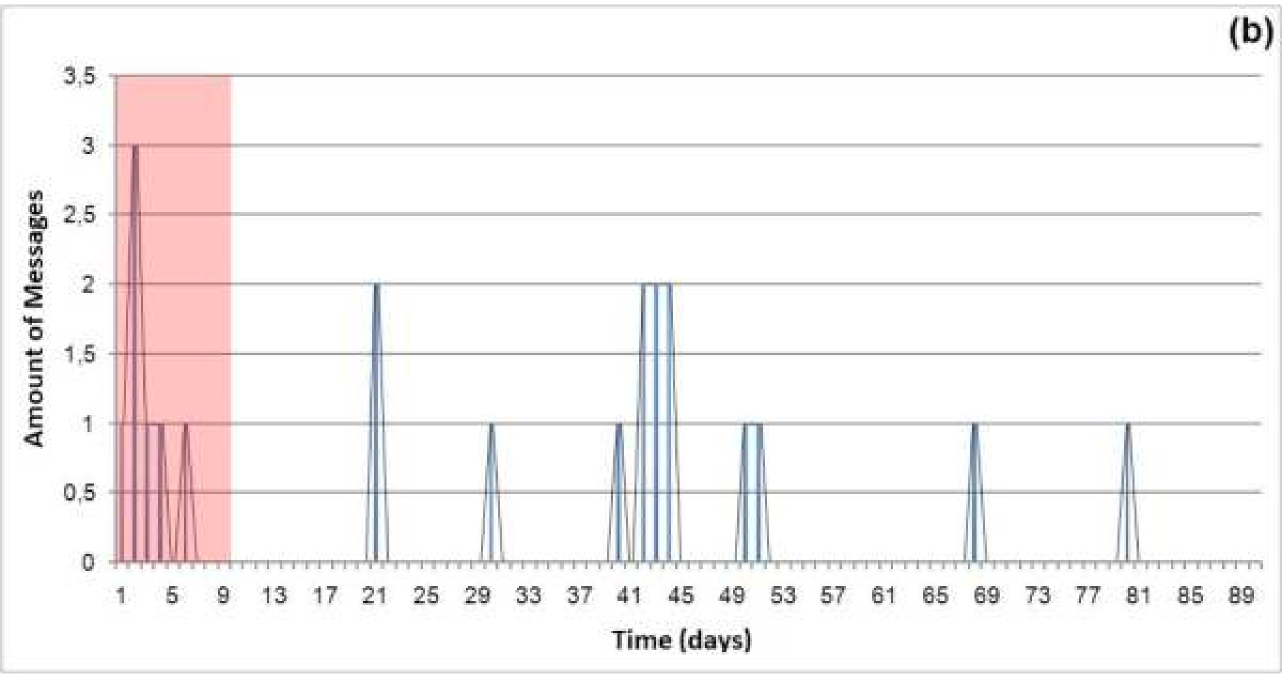


Figure 5: (a) *fepessoinha's* mentioned tweets; (b) *fepessoinha's* private messages.

6.2. Scenario 2

The first part of Figure 6 (the red colored part before nine days) compares the strategies to acquire followers used by *fepessoinha* and *scarina*. We can see that although *scarina* gained influence more slowly than *fepessoinha*, both bots have very close values of Klout score after nine days. This suggests the action of keeping only reciprocated follow links is not penalized by Klout. Thus, users can freely collude with others in order to gain influence in the Twitter social network. Indeed, as observed by Ghosh, *et al.* (2012b), even celebrities like Lady Gaga and Barack Obama reciprocate follow links as a way to be polite or even to keep those connections.

After nine days of experiments, at the end of the first process of collecting and excluding users, the process of posting tweets starts, which initiates Scenario 3.

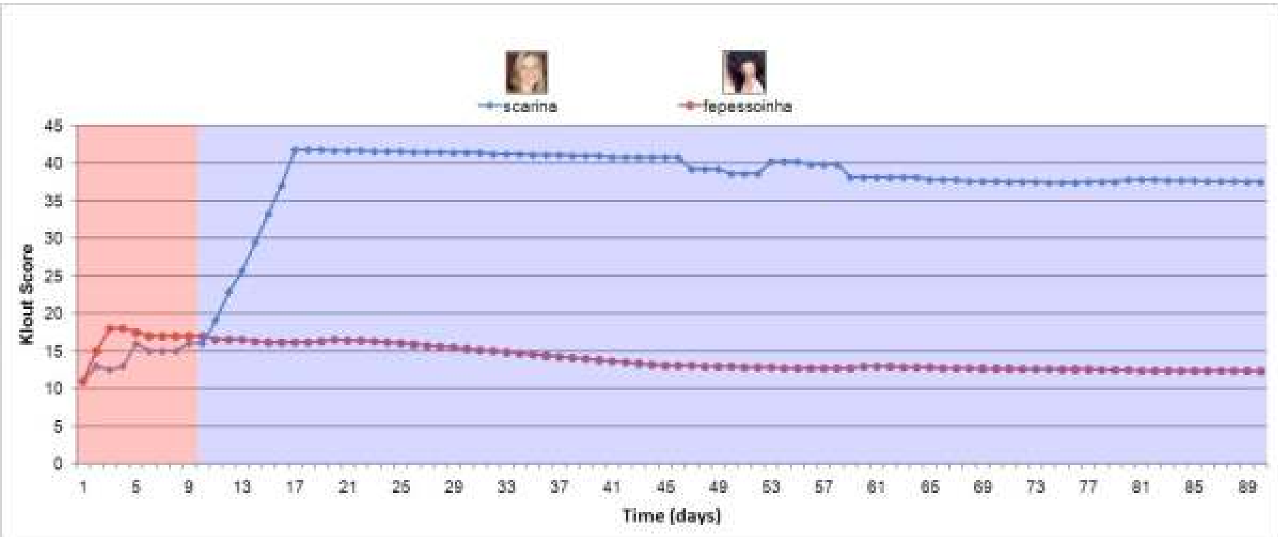


Figure 6: Comparing the Klout Scores of *scarina* and *fepessoinha*.

6.3. Scenario 3

In Figure 6, we can see that from the tenth day, there is a sharp discrepancy between bots *scarina* and *fepessoinha*. The *scarina* bot began to publish tweets from the tenth day, and on day

17 it reaches 41.8 on the Klout Score. We noticed in Scenario 3 that the *scarina* bot began to acquire new followers even without following them first, probably due to the relevance of its posted content. We can conclude from this experiment that to reach a high degree of influence, it is not enough just to follow users, as in Scenario 1. Therefore, the Klout system also takes into account a user's posts in order to increase level of influence. In short, the *scarina* bot arrived at a maximum value of 41.8 in Klout, while the *fepessoinha* bot reached a maximum value of 18.

In addition to the value received in Klout, *scarina* achieved 86 in Twitalyzer. One can see a precise comparison of the highest levels of influence in Figure 7.

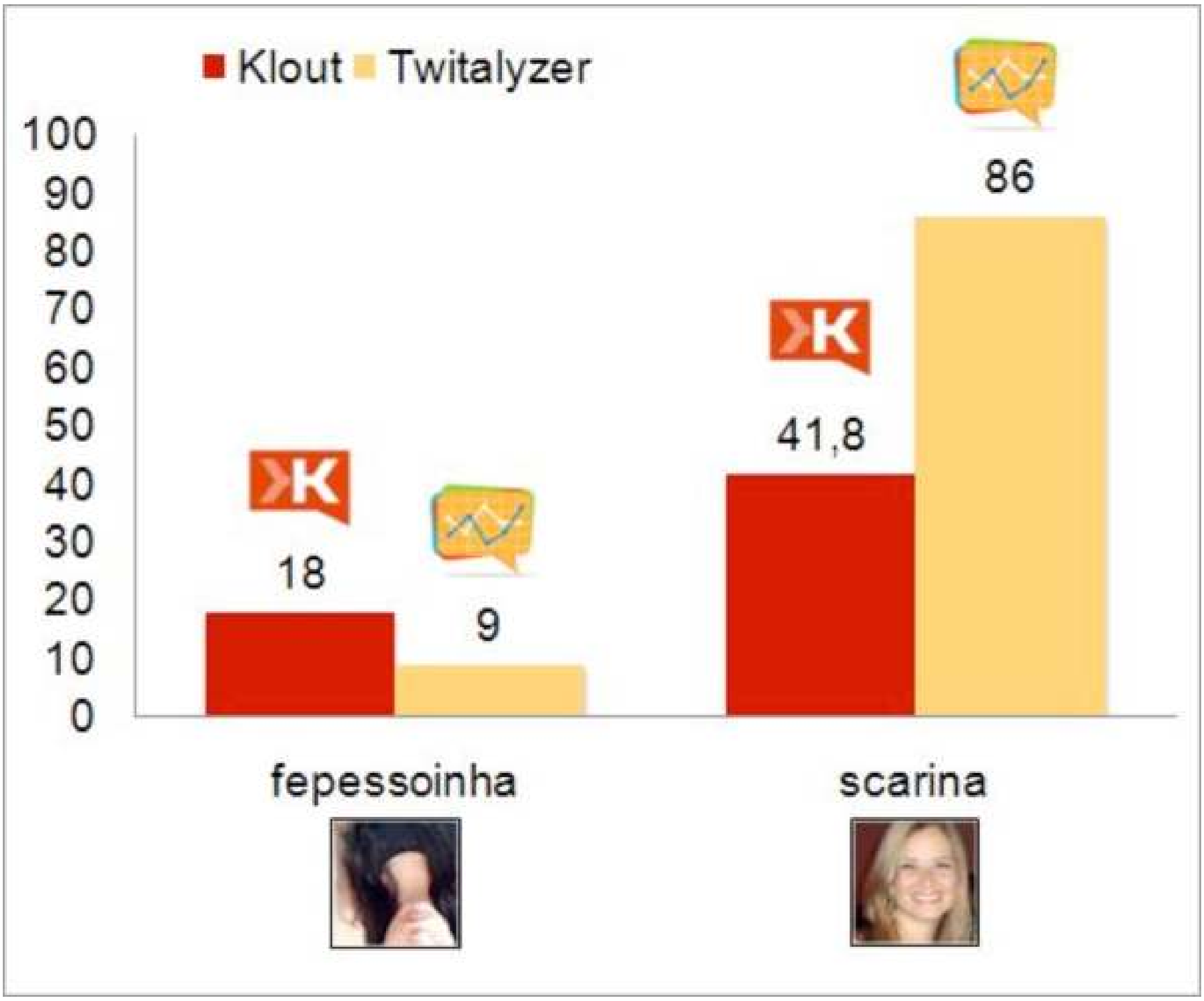


Figure 7: Comparison of Klout and Twitalyzer of *fepessoinha* and *scarina*.

6.4. Scenario 4

The results for Scenario 4 can be clearly seen in Figure 8. We can see big highs and lows corresponding to the interruptions in the execution of the algorithm responsible for the posting of tweets. There were two interruptions, the first one taking place on the 46th day. We kept this algorithm deactivated for six days. We then reinitiated the bot's complete process (following and excluding users as well as posting tweets) and kept it running for six days, before we interrupted it again. At that point, we only restarted the process for posting tweets. We can observe in Figure 8 that the absence of new tweets (with the interruption of the execution), slightly lowered the bot's degree of influence, according to its Klout score.

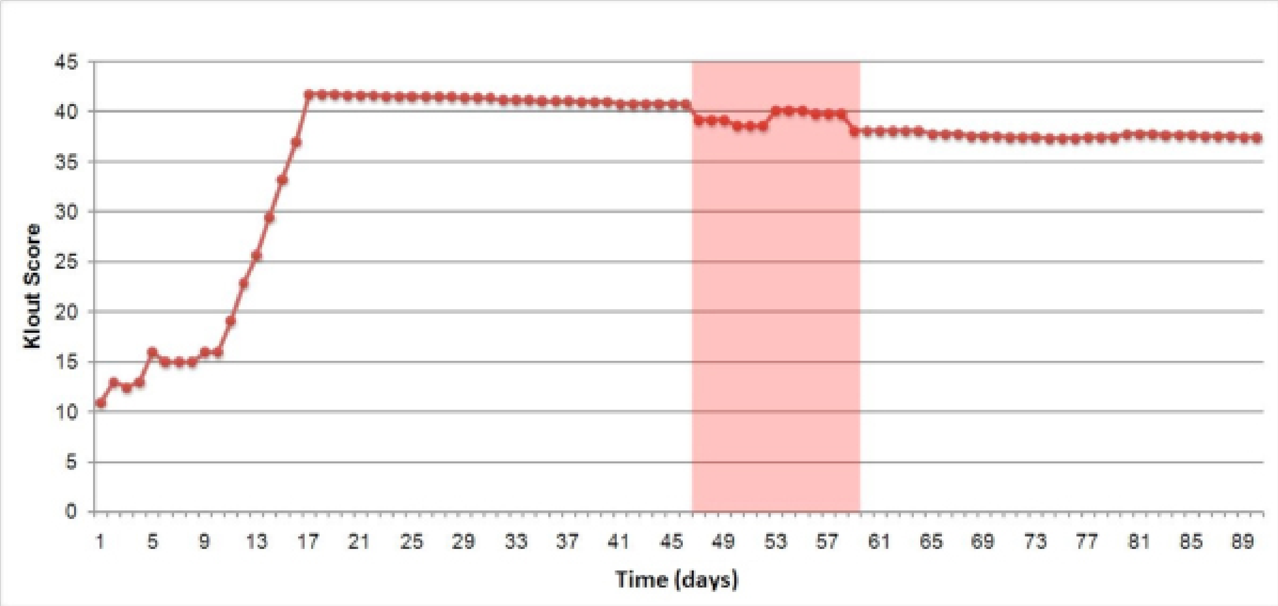
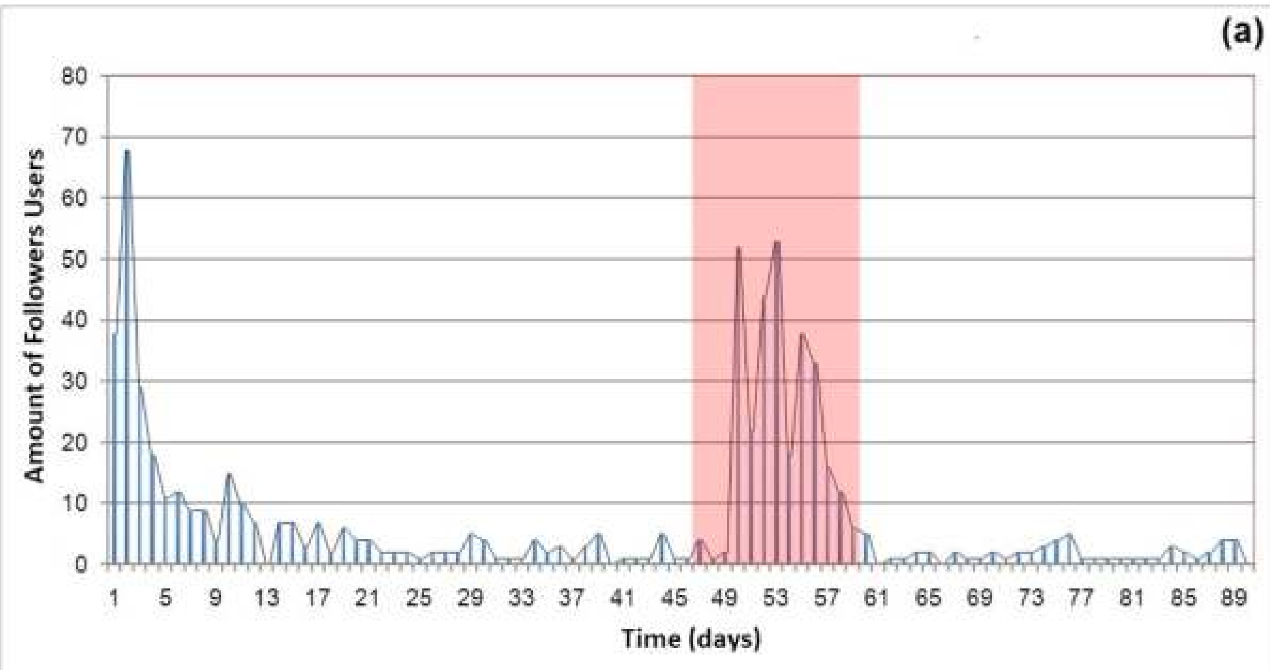


Figure 8: Klout Score results of *scarina* bot.

We can see the same behavior, illustrated in Figure 9 corresponding to the followers, mentions and messages, particularly after the first interruption. At the beginning of the 46th day, during the initial process and after the first interruption, *scarina* attracted new followers. By the end of the 90th day of the experiment, the bot had gained a total of 691 followers.



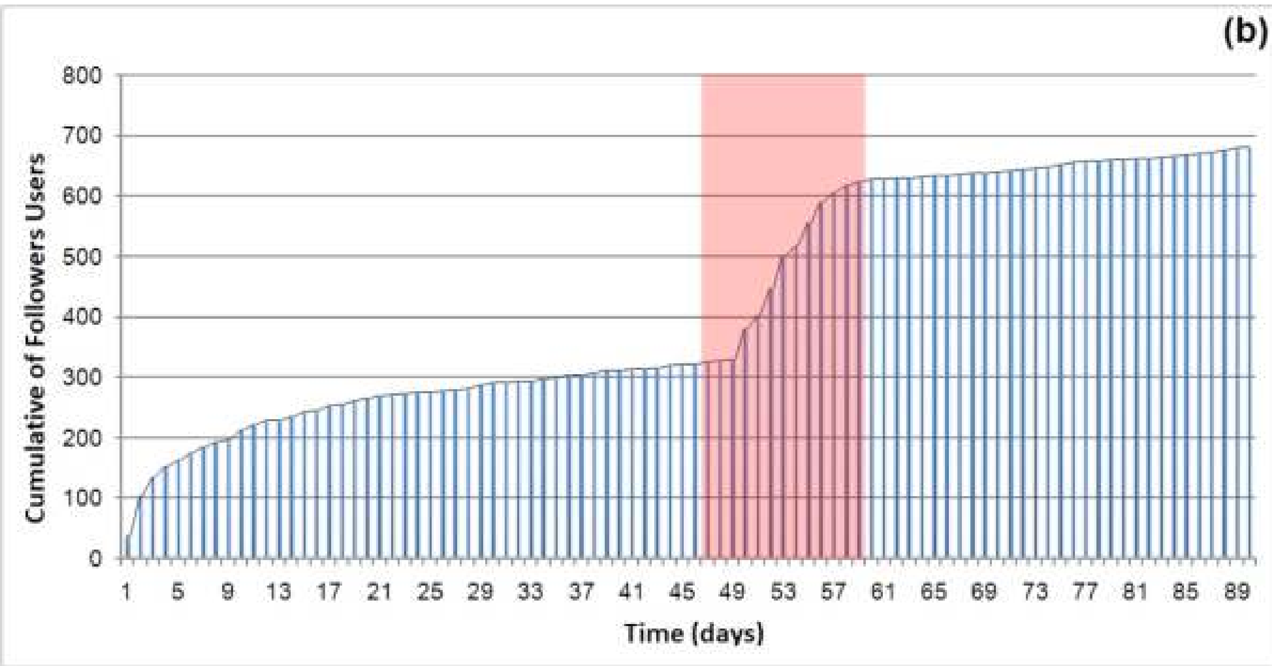
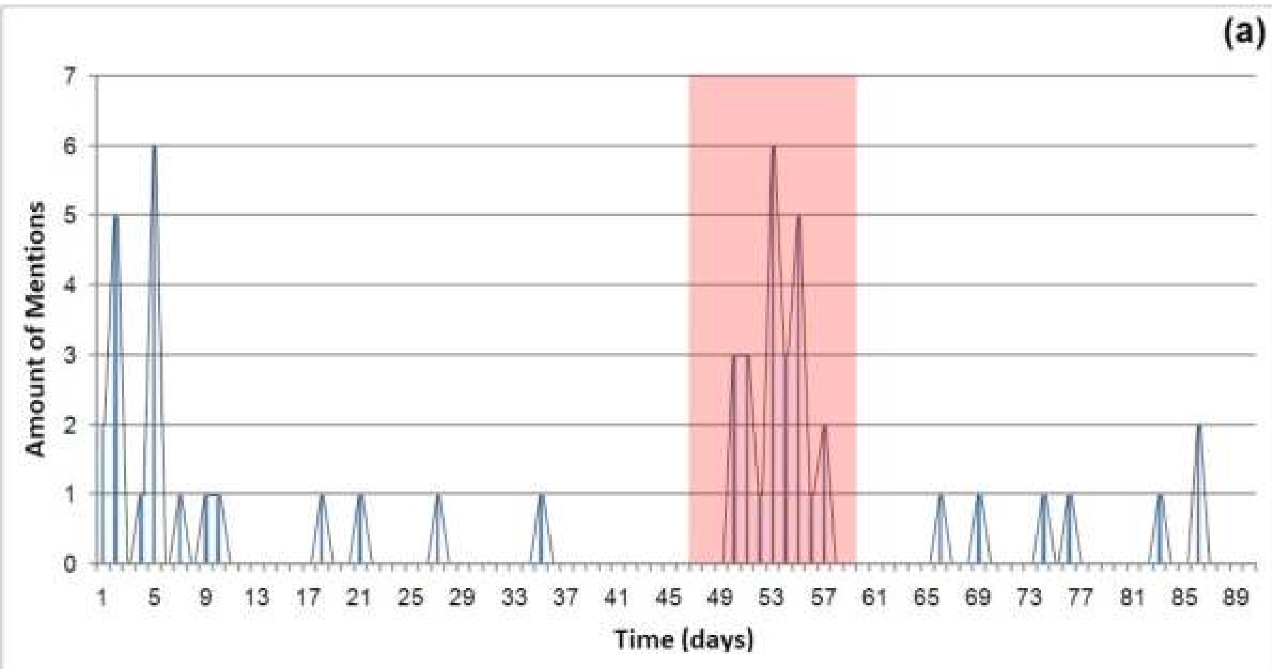


Figure 9: (a) *Scarina's* followers for each day; (b) Cumulative number of *scarina's* followers.

In Figure 10, we can also see that beginning on the 46th day, *scarina* receives new messages and mentions. Because the initial process had been executed again, gaining new followers, the majority of the mentions and messages were questions about the identity of *scarina*, and from where users might know it. In the end of the 90–day experiment, *scarina* had received 52 mentions and messages.



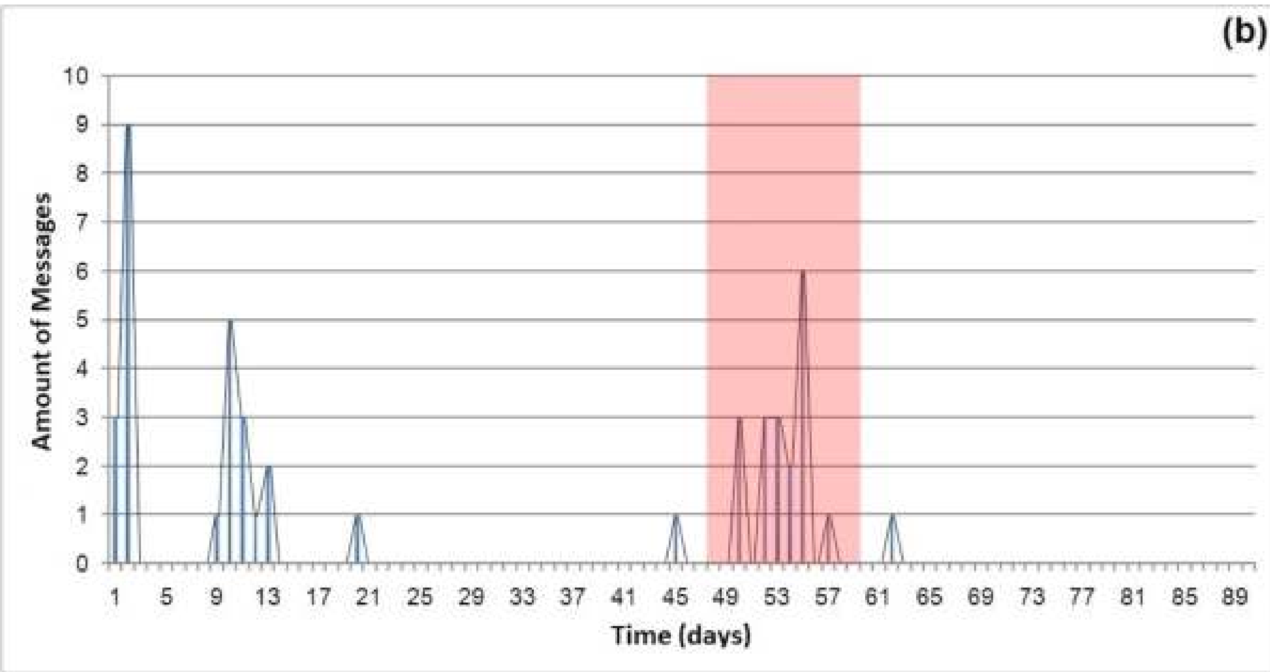


Figure 10: (a) *Scarina's* mentioned tweets per day; (b) *Scarina's* private messages per day.

Scarina posted 4,997 tweets during the 90 days of the experiment. Of this total, only six of the tweets were made favorite by users following the bot. By the end of the 90–day experiment, we had also received a total of 94 retweets and 109 tweet replies.

6.5. Final results

By the end of the 90–day experiment, we had collected values from Klout and Twitalyzer describing the most influential accounts on Twitter. We then completed a final comparison with the *fepessoinha* and *scarina* bots (see Figure 11). The Brazilian comedian Rafinha Bastos, the American rapper Snoop Dogg, U.S. President Barack Obama and the TV presenter Luciano Huck, were amongst the most influential. In addition, we included in this comparison the Twitter accounts of people who are influential in the area of analysis of social networks: Lada Adamic [3] (professor at University of Michigan) and Virgílio A.F. Almeida [4] (a full professor from UFMG). We can see that even though the bots were automatic, they reached similar or higher values of influence in Klout and Twitalyzer, than people of great reputation.

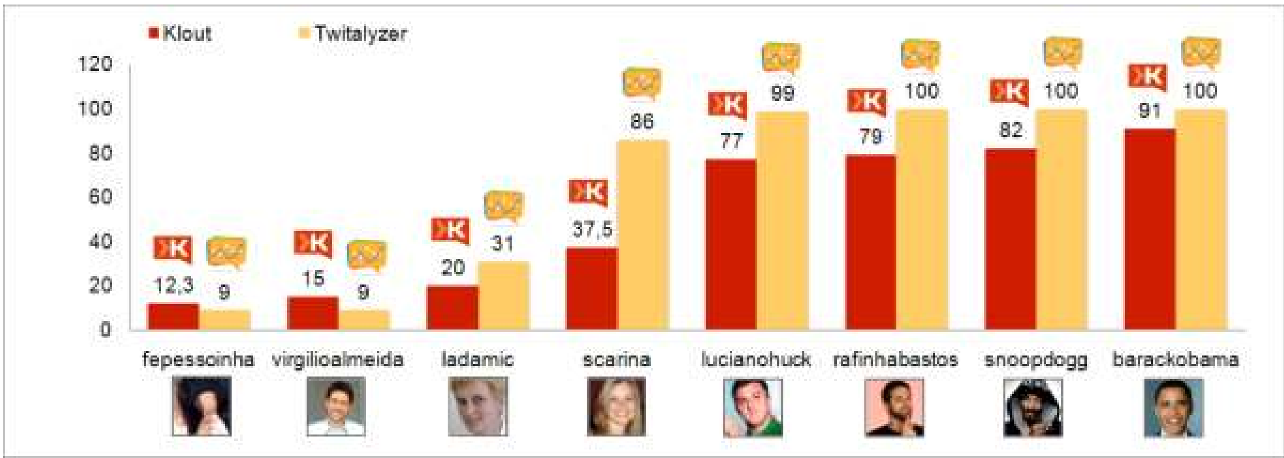


Figure 11: Comparison of Klout and Twitalyzer with other real users.

7. Conclusions and future work

In this paper we described how we created simple robots able of interacting through Twitter accounts as if they were regular users in the network, by exchanging information, following and being followed by users, for a period of 90 days. The strategies used in the bots were very simple, consisting of following only users that followed the bots and posting tweets about popular and focused topics. Even with this automatic and predictive behavior, the bots received significant influence score in two systems that measure influence: Klout and Twitalyzer. One of the bots reached an influence score close to some celebrities and individuals with a high reputation on Twitter.

Our results suggest that many Twitter users need to amass social capital, which is the ideal scenario for bots to infiltrate the system. The users who followed our bots might not be interested in the content they post. They might be interested in reciprocating the follow link as an attempt to keep our bots as part of their audiences. Thus, our results not only highlight vulnerabilities of popular two influential score systems, but they also reveal aspects of the dynamics of user follower/followee behavior in Twitter.

There are some interesting future directions we would like to pursue next. First, we would like to explore a large number of scenarios and bots behaviors, as attempts to quantify which strategies work better to attract new followers and retweets. Second, we would like to propose and evaluate influence ranking mechanisms that less susceptible to robots or automatic activity in Twitter.

About the authors

Johnnatan Messias is an undergraduate in computer science at the Federal University of Ouro Preto (UFOP, Universidade Federal de Ouro Preto), Brazil.
Web: http://lattes.cnpq.br/9861859712648217
E–mail: johnnatan20 [at] gmail [dot] com

Lucas Schmidt is a Ms.C. student in computer science at the Federal University of Ouro Preto, Brazil.
E–mail: lucasscf [at] gmail [dot] com

Ricardo Augusto Rabelo Oliveira is an assistant professor and department chief of the Computer Science Department at Federal University of Ouro Preto, Brazil. He obtained his Ph.D. in 2008 at the Federal University of Minas Gerais (UFMG, Universidade Federal de Minas Gerais). His research plans are focused on mobile and network systems.
Web: http://www.decom.ufop.br/imobilis
E–mail: rrabelo [at] gmail [dot] com

Fabrício Benevenuto is an assistant professor in the Computer Science Department at the Federal University of Minas Gerais (UFMG). His Ph.D. dissertation won the CAPES thesis award as the best Brazilian thesis in computer science defended in 2010 and he was recently elected a member of the Brazilian Academy of Sciences.
Web: http://www.dcc.ufmg.br/~fabricio
E–mail: fabricio [at] dcc [dot] ufmg [dot] br

Acknowledgments

This work was partially funded by InWeb — The Brazilian National Institute of Science and Technology for the Web (grant MCT/CNPq 573871/2008–6), and by the authors' individual grants from CNPq, CAPES e FAPEMIG.

Notes

1. The photos used in the profiles of both bots belong to acquaintances of the authors who authorized their use in bots.

2. See http://www.globo.com.

3. https://twitter.com/ladamic.

4. https://twitter.com/virgilioalmeida.

References

I. Anger and C. Kittl, 2011. "Measuring influence on Twitter," *i–KNOW '11: Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies*, article number 31.

E. Bakshy, J.M. Hofman, W.A. Mason, and D.J. Watts, 2011. "Everyone's an influencer:

Quantifying influence on Twitter," *WSDM '11: Proceedings of the fourth ACM International Conference on Web Search and Data Mining*, pp. 65–74.

P.E. Brown and J. Feng, 2011. "Measuring user influence on Twitter using modified k–shell decomposition," *Fifth International AAAI Conference on Weblogs and Social Media*, at http://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/view/3843, accessed 18 June 2013.

M. Cha, H. Haddadi, F. Benevenuto, and K.P. Gummadi, 2010. "Measuring user influence in Twitter: The million follower fallacy," *ICWSM '10: Proceedings of the International AAAI Conference on Weblogs and Socia Media*, version at http://snap.stanford.edu/class/cs224w-readings/cha10influence.pdf, accessed 18 June 2013.

M. Gladwell, 2000. *The tipping point: How little things can make a big difference*. Boston: Little, Brown.

S. Ghosh, N. Sharma, F. Benevenuto, N. Ganguly, and K. Gummadi, 2012a. "Cognos: Crowdsourcing search for topic experts in microblogs," *SIGIR '12: Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 575–590.

S. Ghosh, B. Viswanath, F. Kooti, N.K. Sharma, G. Korlam, F. Benevenuto, N. Ganguly, and K.P. Gummadi, 2012b. "Understanding and combating link farming in the Twitter social network," *WWW '12: Proceedings of the 21st International Conference on World Wide Web,* pp. 61–70.

E. Katz and P.F. Lazarsfeld, 2006. *Personal influence: The part played by people in the flow of mass communications*. Second edition. New Brunswick, N.J.: Transaction Publishers.

E. Keller and J. Berry, 2003. *The influentials*. New York: Free Press.

M. Kirkpatrick, 2011. "There are now 155m tweets posted per day, triple the number a year ago" (6 April), at http://readwrite.com/2011/04/06/there_are_now_155m_tweets_posted_per_day_triple_th, accessed 19 June 2013.

C. Lee, H. Kwak, H. Park, and S. Moon, 2010. "Finding influentials based on the temporal order of information adoption in Twitter," *WWW '10: Proceedings of the 19th International Conference on World Wide Web*, pp. 1,137–1,138.

D. Leonhardt, 2011. "A better way to measure Twitter influence," *New York Times* (24 March), at http://6thfloor.blogs.nytimes.com/2011/03/24/a-better-way-to-measure-twitter-influence/, accessed 19 June 2013.

A. Pal and S. Counts, 2011. "Identifying topical authorities in microblogs," *WSDM '11: Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, pp. 45–54.

H. Purohit, Y. Ruan, A. Joshi, S. Parthasaranthy, and A. Sheth, 2011. "Understanding user–community engagement by multi–faceted features: A case study on Twitter," *WWW 2011 Workshop on Social Media Engagement (SoME)*, at http://www.cse.ohio-state.edu/~ruan/papers/purohit_some11.pdf, accessed 19 June 2013.

E.M. Rogers, 2003. *Diffusion of innovations*. Fifth edition. New York: Free Press.

D.M. Romero, W. Galuba, S. Asur, and B.A. Huberman, 2011. "Influence and passivity in social media," *WWW '11: Proceedings of the 20th International Conference Companion on World Wide Web*, pp. 113–114.

J. Weng, E.–P. Lim, J. Jiang, and Q. He, 2010. "TwitterRank: Finding topic–sensitive influential twitterers," *WSDM '10: Proceedings of the Third ACM International Conference on Web Search and Data Mining*, pp. 261–270.

J.L.S. Yan and E. Kaziunas, 2012. "What is a tweet worth? Measuring the value of social media for an academic institution," *iConference '12: Proceedings of the 2012 iConference*, pp. 565–566.

Editorial history

Received 23 August 2012; revised 10 April 2013; revised 13 June 2013; accepted 14 June 2013.

You followed my bot! Transforming robots into influential users in Twitter
by Johnnatan Messias, Lucas Schmidt, Ricardo Oliveira, and Fabrício Benevenuto.
First Monday, Volume 18, Number 7 - 1 July 2013
http://firstmonday.org/ojs/index.php/fm/rt/printerFriendly/4217/3700

doi:10.5210/fm.v18i7.4217