A Close Look at Tinder Bots

Tahora H. Nazer* Fred Morstatter* Gareth Tyson† Huan Liu*

Abstract

Tinder is a popular dating app that allows users to discover potential dating partners with close geographical proximity. Tinder is the first dating app in several countries and has more than 50 million users. However, many of these users are bots with malicious intent. The first step in dealing with this issue is understanding the characteristics of Tinder bots. Toward this aim, we have proposed a ground truth collection method to acquire bots to study. Our method combines honeypot methods and manual annotation. We find that probing messages is a reliable method to distinguish bots from humans as bots promote malicious URLs and direct users to phishing sites. Our observations on the collected bots show that they are more complex than bots that are studied in other social media sites. Tinder bots have profiles that are very hard to differentiate from normal users. We explore activity and profiles of these bots and report the characteristics that can be used in building a supervised learning approach for bot detection.

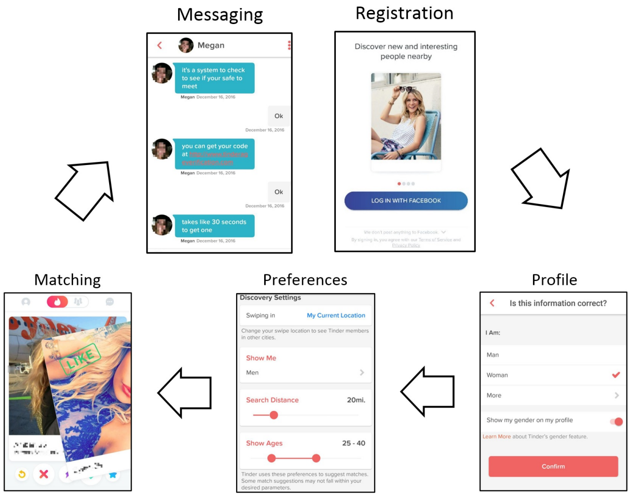
1 Introduction

Tinder[1] is one of the most popular dating applications for Android and iOS mobile phones. Tinder is recognized as the most downloaded app in 18 countries with the biggest app markets. In several countries including USA, UK, Canada, and Australia, Tinder is the most popular dating app[2]. Tinder has more than 50 million users [1] and they spend around 77 minutes on it [5] every day.

Tinder users look for dating partners with the assumption that the profile they see belongs to a real person (see Section 2). However, bots are responsible for 51.8% of web traffic [12] and Tinder is not an exception. Bots on Tinder interfere activities of normal users. They advertise inappropriate or phishing websites, promote subscription to unwanted services, and encourage users to switch to other platforms.

To prevent these malicious activities, Tinder allows users to report bots. Based on our observations, fake accounts and inappropriate messaging results in the account being banned in less than 24 hours. However, some of the Tinder bots are very complicated and disguise themselves well among humans. Therefore, bot detection methods are required that can distinguish these bots.

To build a model that automatically detects bots, we need to extract discriminating features of bots. The first challenge is collecting a sample of bots. On Tinder, there is no way to access user profiles unless they are shown by an unknown heuristics as a potential dating parter who is in a close geographical proximity. This causes data collection to be bound by the number of users that are suggested by Tinder and the location being monitored. Also, previous ground truth collection methods cannot be applied (Section 3). To overcome this challenge we have proposed a method based on a honeypot method with manual intervention (Section 4). Using this method we collected 146 bots.



Megan

it's a system to check to see if your safe to meet
Megan December 16, 2016

Ok
December 16, 2016

you can get your code at http://www.tinderag everification.com
Megan December 16, 2016

Ok
December 16, 2016

takes like 30 seconds to get one
Megan December 16, 2016

Discover new and interesting people nearby

LOG IN WITH FACEBOOK

We don't post anything to Facebook.
By signing in, you agree with our Terms of Service and Privacy Policy

Is this information correct?

I Am:
Man
Woman ✓
More
Show my gender on my profile ●
Learn More about Tinder's gender feature.
Confirm

Discovery Settings
Swiping in: My Current Location
Change your swipe location to see Tinder members in other cities.
Show Me: Men
Search Distance: 20mi.
Show Ages: 25 - 40
Tinder uses these preferences to suggest matches. Some match suggestions may not fall within your desired parameters.

Figure 1: User life-cycle on Tinder. After a user joins Tinder, majority of the time is spent for matching and messaging.

The second challenge is extracting features that discriminate bots from humans. We studied profiles and messaging behaviors of bots. As discussed in Section 5, majority of bot profiles include a bio, school, and job. Moreover, they have an average of 60 connections on Facebook. Profile images are original in 71% of time. However, their accounts are not linked to an Instagram or Spotify.

We have collected 570 messages originated from a bot. These messages contain large number of URLs which direct users to malicious/phishing websites or unwanted services. They use URL shortening services and it disables users to judge the content of the web pages based on the URL. On average, conversations with bots are 6-message long and each message has 15 words. Tinder bots that we observed are not intelligent in their messages and their responses does not change based on the messages that they receive.

*Arizona State University
{tahora.nazer,fred.morstatter,huan.liu}@asu.edu
†Queen Mary University of London
g.tyson@qmul.ac.uk

[1] https://www.gotinder.com/

[2] https://goo.gl/JP0RV5

2 Overview of Tinder

The life cycle of a user on Tinder starts with registration (see Figure 1). Users can log in using their Facebook credentials without creating a new profile separately for Tinder. The first time that user signs in to the system, a Tinder account will be automatically created and the user will be able to manage the profile and perform other tasks.

Tinder profiles have fixed and editable parts. Name and age are automatically obtained from users' Facebook profile and cannot be changed in Tinder. This information is public and will be visible to potential matches. Gender is also retrieved from Facebook, however, users can modify it after the first log in. Moreover, they can choose to hide or show their gender. Users can import their Facebook images, upload new photos, write a short bio (up to 500 characters), and import work and school from Facebook. They can also link their Instagram and Spotify accounts to Tinder.

In the next step, users choose their matching preferences. Gender, age, and distance are three criteria that can be modified. Gender of users who are suggested as a potential match is either "Men" or "Women". Age can be 18 or higher (all ages over 55 are grouped as 55+). Tinder suggest users who are 1 to 100 miles away.

Potential matches in a close geographical proximity are sorted and shown to users based on an unknown heuristic. Users can either "Super Like"(the action of expressing great interest), "Like" (the action of expressing interest or swiping right), "Pass" (the action of showing reluctance or swiping left) them. A hundred likes are allowed every 24 hours. In the paid version of Tinder, "Tinder Plus", uses have unlimited number of matches.

If two users both swipe right on each other, they will match and they can exchange messages. The conversation is considered successful if it results in the exchange of contact information. Users can "Unmatch" at any time after a match has happened. The unmatched users will be inaccessible afterward and their information and messages will be removed from user's activity page.

3 Ground Truth Collection Methods

There are at least three ground truth acquisition methods for bot detection on social media: user status, honeypots, and manual annotation. Users can be suspended on social media sites to maintain a safe environment. The suspension is a result of violating rules of the social media platform. Examples of harmful activities that can cause a suspension on Twitter are abusive content such as violating copyrights, abusive behavior such as creating multiple accounts or sharing private information, or spam. Malicious bots actively perform aforementioned activities to disturb activities of normal users. Hence, it is a reasonable assumption to consider suspended users as bots.

Honeypots are a more recent way of gathering reliable ground truth with minimum manual interference [6, 7]. They are bots that do not have human-like behaviors and do not intervene with normal activities of human users. Honeypots are designed to be easy for humans to distinguish. So, all the users that interact with bots are expected to be bots. However, there are always real users who interact with honeypots for sake of acquiring more visibility.

In manual annotation, users are inspected by human annotators and then labeled as human or bot. In the labeling process, annotators examine different features of an account such as profile picture, bio, account age, posts, and posting behaviors. This method is the most common [2, 3] but it is time-consuming, expensive, and error-prone.

4 Collecting Tinder Bots

We have used a combination of a honeypot method and manual annotation for collecting ground truth. We have 3 female and 3 male honeypot accounts on Tinder. Each account has 3 profile images and a short bio. Their names are chosen from common names in the United States and their location is Phoenix, Arizona. They have been active for three months and one of them was a "pro" account for 2 months. They like as many dating partners as Tinder allows.

The honeypots monitor matches and reply to messages that they receive (they never start a conversation). To avoid any intervention with activities of humans, honeypots stop the interaction if any of the conditions bellow is satisfied:

- The sender stops messaging the honeypot before we identify user as bot or human.

- The sender posts a personalized message which contains an information which can only be interpreted by humans. For example, they mention an object in the profile photo. Such users are considered as humans.

- The sender start the conversation with a phone number, invites the honeypot to move the discussion to another platform such as Skype, or sends a URL and encourages the honeypot to visit it. In any of these cases, we consider the corresponding user as a bot.

This method is automatic in liking potential partners, responding to messages, collecting profile and messages of matches, extracting URLs from messages. Deciding if a conversation is controlled by a bot, investigating URLs and judging content of the websites has been performed manually. This process is time-consuming and needs human supervision. However, based on the limitation exposed by Tinder, we found it a reliable method by which we could collect a ground truth of bots. Using the method above in the course of 3 months we have collected 146 bots among 623 users that communicated with our honeypots.

5 Scrutinizing Tinder Bots

We collected 146 bots in total using the method explained in Section 4. We use this dataset [3] to study characteristics of bots in terms of their profiles (bio, image, and connections) and messaging behaviors. The ultimate goal is building a bot detection mechanism for Tinder. As seen in previous studies [7, 6], supervised mechanisms require larger ground truth datasets. Hence, the collected dataset is only used for a proof of concept showing how Tinder bots are different from Tinder users and other bots (e.g. Twitter bots).

5.1 Profile

As shown in Table 1, profiles of Tinder bots are very close to profiles of humans: 59% have a bio in their profile with an average length of 18 words or 104 characters, 68% mention a school that they attended, and 32% list a job. Only profile features that are missing from bots' profiles (except for one bot) is Spotify and Instagram accounts.

Table 1: Profile characteristics of Tinder bots

Feature	Value
Total bots with bio	86
Total bots with job	48
Total bots with school	99
Total bots with Spotify account	0
Total bots with Instagram account	1
Total bots with copied profile photo	42
Avg characters in bio	103.91
Avg words in bio	17.78
Avg Facebook connection	60.43

Inspecting connections of bots reveals that they have an average of 60 friends on their Facebook account. As shown in Figure 2, bots have up to 302 connections with a concentration between 60 to 90 friends.

Bots have 0 to 4 photos in their profile and the majority have 4 photos as shown in Figure 3. We also used Google reverse image search on these photos and 42 bots have at least one photo which is copied from the web. All the aforementioned properties show that these bots have complex profiles and it is hard to be distinguished.

5.2 Activities

The activity by which we distinguished bots from humans is messaging. The messaging behavior of bots shows different characteristics in comparison to the average behavior of users on Tinder. Previous studies [13] reported average length of conversation to be 9 messages, however, our bots participated in conversations with an average of 6 messages. The average length of messages from bots is 75 characters or 15 words. In contrast, the average length of messages for normal users on Tinder is 59 characters or 11 words [13].

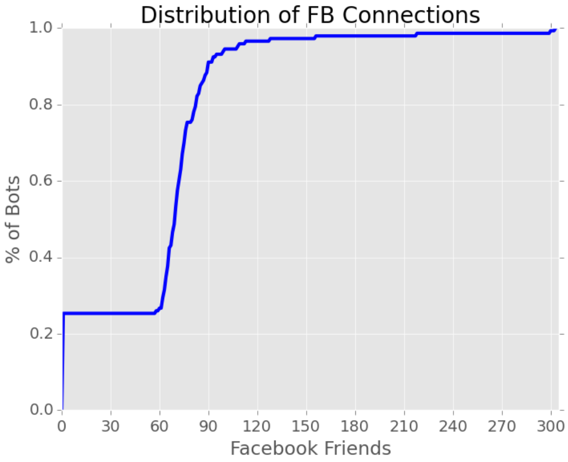


Figure 2: CDF of bots' Facebook connections.

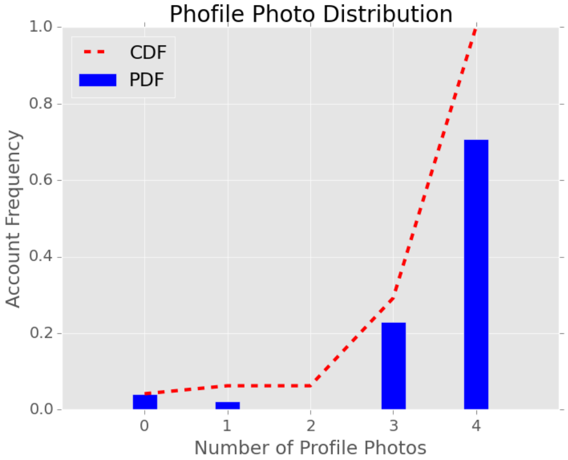


Figure 3: Distribution of profile photos.

Regarding the content of messages, we witnessed a large number of URLs in the messages sent from bots. Many of these URLs are repeated which can be a sign of collaboration between bots; i.e. having a common controller which makes them more effective [9]. The majority of these URLs are shortened with "bit.ly" service. Using shortened URLs helps with trapping more victims as the user cannot have any context before following the link.

6 Related Work

Studies on Tinder since its foundation in 2012 has been mostly focused on social aspects of it [4, 8, 11]. One study finds six motivations for using Tinder which vary based on age and gender [8]. The authors argue that the first motivation of using Tinder is love and Tinder is a "new way of committed romantic relationship".

In another study [4], the authors analyze the negative

[3] Data collection code is available at `https://goo.gl/o6JqOC`

Table 2: Messaging behaviors of Tinder bots

Feature	Value
Total number of messages from bots	570
Messages with a URL	84
Unique URLs	23
URLs shortened using “bit.ly”	54
Avg characters per message	75.39
Avg words per message	14.9
Avg length of conversations	6.12

behavior of users, “crude performance of masculinity”. This work which is based on messages and comments on Tinder Nightmares Instagram page [4], explains how users response to abusive messages and how responses diffuse.

Quantitative analysis of user behaviors on Tinder [10] shows the difference between male and female users. Women gain more matches, their messages are longer, and there is a longer lapse between the match and their first message. Messaging behaviors show that conversations are mainly initiated by men and successful matches result in an exchange of phone numbers in the first 20 messages [13].

The bots that are created and controlled by users are often the target of bot detection methods. However, there are cases that the service provider uses bots to increase the number of subscribers, change female-to-male ratio on its site, and lure users to subscribe.

7 Conclusion

We proposed a honeypot method with some level of manual intervention for finding bots on Tinder. We collected 146 bots and our studies showed that they are more complex than previously observed bots on other social media sites. They are very successful in disguising among normal users: they have 60 Facebook friends on average, use real photos, and have complete profiles (with bio, school, and job) . These properties make the bot detection task very challenging.

Tinder bots, on the other hand, share a large number of malicious URLs. This was the main feature we used to distinguish bots from normal users. Extraction and investigation of URLs can be automated to some extent. However, it can be applied after an interaction between bots and benign users. This can cause disturbance and inconvenience.

8 Future Directions

The ultimate goal is having effective methods for detecting bots on Tinder. Our proposed method is successful in detecting bots after they interact with users. However, detecting bots before messaging users is more advantageous. We believe that further examination of bot profiles, if performed in large scale, will reveal characteristics that can be used in supervised learning methods.

9 Acknowledgments

This research was supported, in part, by Office of Naval Research grant N000141310835. The information reported here does not reflect the position or the policy of the funding agencies. The authors would like to thank Zafar Gilani from University of Cambridge for his valuable feedback.

References

[1] N. Bilton, *Tinder, the fast-growing dating app, taps an age-old truth*. `https://goo.gl/uKHrR0`, Oct. 29, 2014.

[2] Z. Chu, S. Gianvecchio, H. Wang, and S. Jajodia, *Who is tweeting on twitter: human, bot, or cyborg?*, in Proceedings of the 26th annual computer security applications conference, ACM, 2010, pp. 21–30.

[3] D. M. Cook, B. Waugh, M. Abdipanah, O. Hashemi, and S. A. Rahman, *Twitter deception and influence: Issues of identity, slacktivism, and puppetry*, Journal of Information Warfare, 13 (2014).

[4] A. Hess and C. Flores, *Simply more than swiping left: A critical analysis of toxic masculine performances on tinder nightmares*, new media & society, (2016), p. 1461444816681540.

[5] A. Kleinman, *The typical tinder user spends 77 minutes tinding every day*. `https://goo.gl/TQ1cyJ`, Nov. 04, 2014.

[6] K. Lee, B. D. Eoff, and J. Caverlee, *Seven months with the devils: A long-term study of content polluters on twitter.*, in ICWSM, 2011.

[7] F. Morstatter, L. Wu, T. H. Nazer, K. M. Carley, and H. Liu, *A new approach to bot detection: striking the balance between precision and recall*, in Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2016.

[8] S. R. Sumter, L. Vandenbosch, and L. Ligtenberg, *Love me tinder: Untangling emerging adults motivations for using the dating application tinder*, Telematics and Informatics, 34 (2017), pp. 67–78.

[9] K. Thomas, C. Grier, D. Song, and V. Paxson, *Suspended accounts in retrospect: an analysis of twitter spam*, in Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference, ACM, 2011, pp. 243–258.

[10] G. Tyson, V. C. Perta, H. Haddadi, and M. C. Seto, *A first look at user activity on tinder*, in Advances in Social Networks Analysis and Mining (ASONAM), 2016 IEEE/ACM International Conference on, IEEE, 2016, pp. 461–466.

[11] J. Ward, *What are you doing on tinder? impression management on a matchmaking mobile app*, Information, Communication & Society, (2016), pp. 1–16.

[12] I. Zeifman, *Bot traffic report 2016*. `https://goo.gl/Go4BQ2`, Jan. 24, 2017.

[13] J. Zhang and T. Yasseri, *What happens after you both swipe right: A statistical description of mobile dating communications*, arXiv preprint arXiv:1607.03320, (2016).

[4] `https://goo.gl/ZJiOO7`