Detecting Like Farm accounts on Facebook

authors' names

Abstract

With the growing popularity of online social networks, reputation manipulation on online social networks has also emerged. A large body of research focus on detection of such activities and OSN accounts used in these activities. However, previous research has shown that fraudsters evade these detection algorithms by mimicking normal users behavior. In this study, we focus on one such problem where Facebook detection algorithm was unable to identify like farm accounts. We plan to identify timeline based features of a user to distinguish a like farm user from a normal user. Finally, we will build a machine learning based classifier to detect like farm users with high recall and low false positive rate.

1 Introduction

The increasing popularity of online social setworks (OSNs) gauged attention of businesses and public figures.Online social networks offer various methods and techniques for promotion and marketing. Facebook page is one such method to assist businesses and public figures to reach out to their audience. Each like on a Facebook page have some monetary value associated to it. According to Blackbaud [1], a like on Facebook is worth approximately 214$.

Facebook page owners are always interested in increasing their Facebook page likes. There exists legitimate ways of promoting Facebook pages such as Facebook ad campaigns. However, there exists a growing market of fake Facebook likes, also known as like farms, which provide Facebook likes much faster and is also cheap [4]. Emiliano et. al [4] created thirteen honeypot Facebook pages and bought likes from different like farms to understand their social characteristics. Authors categorized like farms into two categories i.e., naive like farms and stealthy like farms. Naive like farms like Facebook pages with a large number of accounts in a short interval of time. However, stealthy like farms mimick normal user's behavior and provide likes in a longer time interval to circumvent anomalous behavior detection algorithms such as CopyCatch [2]. In addition to that, they discovered that only a few likes generated by like farms and a limited number of Facebook account were removed. Thus, we need an algorithm to detect the activities and correctly classify the accounts of like farms.

In this study, we will develop a machine learning based approach to classify like farm accounts. To this end, we will identify the distinguishing features of a user's timeline to build multiple classifiers to achieve best accuracy. This problem is challenging because our classifier should produce high recall and low false positive rates. Our objective is to identify the usefulness of timeline based features in the identification of like farm accounts.

2 Related Work

There is a large body of research on detection of anomalous activities and fake accounts in online social networks. However, we specifically focus on the detection of fake accounts used by like farms to increase page likes.

Lee et al. [5] trained a machine learning classifier based on content features, social graph characteristics, and posting patterns to detect spammers on MySpace and Twitter. Stringhini et al. [7] trained a machine learning classifier based on features such as message similarity and URL ratio to detect spammers on Facebook and Twitter. Boshmaf et al. [3] detected fake accounts on Facebook and Tuenti by predicting targets based on features such as gender, number of friends, time since last update, etc. Song et al. [6] trained a machine learning classifier to detect reputation manipulation targets on Twitter based on features such as retweet time distribution, ratio of the most dominant application, number of unreachable retweeters, and number of received clicks.

Beutel et al. [2] proposed a bipartite graph clustering algorithm to detect suspicious lockstep activity patterns on Facebook, also known as CopyCatch, which is currently deployed on Facebook. Emiliano et al. [4] used honeypot pages and bought likes from like farms. They categorized the like farms into two categories i.e., stealthy like farms and naive like farms. While naive like farms gives likes in bursts and are easy to detect, stealthy like farms mimick normal user's behavior and give likes in longer interval of time to evade detection algorithms such as CopyCatch. In this paper, we also specifically focus on the detection of such accounts used by like farms to boost page likes.

3 The Proposed Work

As discussed earlier, the current methodologies are unable to detect activities and accounts of like farm accounts. Our idea is that we should look into the activity patterns of the users that inherently distinguish real user activity from an automated/batch activity. Let's say for example, normal users' Facebook status updates are more unique and original since they corresponds to the life events of an actual user. However, fake users are automated or work in batch mode. It is not possible for fake users to produce unique content. Thus, uniqueness in content sharing of such users may be less than normal users. Our focus in this research work is to identify such distinguishable features or patterns of a users' timeline which can assist us in modeling a classifier to detect like farm users.

3.1 Timeline Features

Some of the features of the timeline of a Facebook user that we plan to explore are:

3.1.1 Types of timeline posts

The posts on users' timeline are very informative to characterize a user. Generally, normal users have more textual posts on their timelines whereas like farm users generally have non-textual posts like, mood statuses, life events etc. We can perform statistical analysis on types of post on user timelines and treat them as input features.

3.1.2 Word Frequency Analysis

The content posted on a timeline of a user can be considered as a document. The word frequency analysis of this document can also be used as a distinguishing feature for normal and like farm users.

3.1.3 Comments and Like frequency

Frequency of comments and like generated by users can be another useful feature that can help distinguish activity pattern between normal users and the like farm users.

3.1.4 Other possible features

What percentage of content generated is original and what percentage of it is just the shared content. The quality of Facebook statues generated by a user, for example, average length of status, percentage of text from dictionary vocabulary and percentage of it having typos, and usage of emojies etc. can also be useful features.

3.2 Page likes

The correlation between the page liking activity of different users can also be useful. Although, this information,stand alone, was not successful in predicting the like farm users but this may help in combination with other features for better identification of like farm users.

Finally, we will use these features to build a classifier to distinguish like farm users from normal users. The primary objective of our classification model is to have high accuracy and low false positives.

4 Plan

Emiliano et al. scrapped the timeline data of the like farm users. They also scrapped data of some normal users as baseline. In addition to timeline data, they scrapped the pages liked by the like farm users and baseline users. We plan to use this data since we have access to it. However, it is important to note that the amount of information crawled for each user depends on the privacy settings of the user. Hence, we can expect some missing values. We also plan to address the challenge of missing values if it will become a hurdle in classification.

Finally for classification, we plan to use a whole range of machine learning algorithms, for example: Artificial Neural Networks, Naive Bayes, Support Vector Machines, Decision Trees and Instance based learning. Furthermore, we plan to employ ensemble methods to see if they can help us to achieve higher accuracy compared to the base classifiers.

In this study, we plan to detect specifically like farm users which are not detected by currently deployed detection algorithms of Facebook such as CopyCatch [2]. Thus, we do not have any concrete baseline results to compare our results.

We would like to acknowledge that this dataset is collected by Emiliano et al. to analyse the shortcomings in CopyCatch and understand the peculiar activity patterns of like farm users. This is an ongoing project.

References

[1] Nonprofits Value a Facebook Like at $214.81. `http://npengage.com/nonprofit-fundraising/nonprofit-value-facebook-like/`.

[2] A. Beutel, W. Xu, Wenkatesan, Chirstopher, and Christos. CopyCatch: Stopping Group Attacks by Spotting Lockstep Behavior in Social Networks. In *WWW*, 2013.

[3] Y. Boshmaf, D. Logothetis, G. Siganos, J. Leria, J. Lorenzo, M. Ripeanu, and K. Beznosov. Integro: Leveraging Victim Prediction for Robust Fake Account Detection in OSNs. In *Network and Distributed System Security Symposium (NDSS)*, 2015.

[4] E. D. Cristofaro, A. Friedman, G. Jourjon, M. A. Kaafar, and M. Z. Shafiq. Paying for Likes?: Understanding Facebook Like Fraud Using Honeypots. In *ACM Internet Measurement Conference (IMC)*, 2014.

[5] K. Lee, J. Caverlee, and S. Webb. Uncovering social spammers: social honeypots + machine learning. In *ACM SIGIR*, 2010.

[6] J. Song, S. Lee, and J. Kim. CrowdTarget: Target-based Detection of Crowdturfing in Online Social Networks. In *ACM CCS*, 2015.

[7] G. Stringhini, C. Kruegel, and G. Vigna. Detecting Spammers on Social Networks. In *Annual Computer Security Applications Conference (ACSAC)*, 2010.