There Are No Free iPads: An Analysis of Survey Scams as a Business

Jason W. Clark
George Mason University

Damon McCoy
George Mason University

Abstract

Spam is a profit-fueled enterprise and cyber-criminals are focusing more of their efforts at growing Online Social Networks, such as Facebook. One of the common methods of monetizing Online Social Network spam is to entice users to click on links promising free gift cards and iPads. However, these links actually lead to ad networks that bombard users with surveys in an attempt to collect personal and contact information that they will sell to other marketers. To date, we lack a solid understanding of this enterprise's full-structure. In this paper, we examined the survey scam process to determine the affiliates/sponsors that are behind this lucrative scam by performing an analysis of five months of Facebook spam data. We provide the first empirical study and analysis of survey scams and demonstrate how to determine which ad networks are sponsoring this spam.

1 Introduction

The growing user bases of Online Social Networking (OSN) sites has become an increasingly lucrative target for profit motivated cyber-criminals, such as the "koobface gang" that targeted Facebook and Twitter with large-scale spam campaigns [7]. These spam campaigns lure users to click on enticing posts, such as "free giveaway" offers for free gift cards and iPads. However, once the user clicks on one of these links they are often instructed to complete a survey prior to receiving their free gift card, iPad or being able to view the advertised video clip. These spamvertised links on Facebook are being monetized by directing users to specialized ad networks that are known as a *Cost Per Action* (CPA) or *lead generation* affiliate based ad networks that pay their affiliates a commission for every "survey" that a user completes.[1] These "surveys" in reality are crafted by clever advertisers and are merely focused at having the user install some profit-generating browser toolbar or rapidly getting the user's contact information so that they can contact them with follow-up offers and finally presenting the user with "limited time discounted" subscriptions to dating sites and magazines. To date, we lack a solid understanding of this enterprise's full-structure.

In this paper, we describe our empirical analysis of spam URLs identified by a popular spam detection Facebook application to better understand this ecosystem and identify which ad networks are involved in sponsoring these Facebook spammers. We find that 73% of the functioning spam URLs are monetized via survey scams and of the 129 unique spam URLs over 50% of these URLs were traced back to four ad networks: Amung.us, CPAlead, ClickBanner, and LifeStreet Media.Based on this analysis, we demonstrated how to extract an affiliate's ID by visiting and interacting with the spam URL. We also develop a *carbon dating* method that allows us to estimate how old an affiliate's account is to understand how proactive the ad networks are at blocking accounts used by spammers. Finally, we joined a number of these ad networks to understand their commission payment structure and expected revenue per click. Collectively, we gain a better understanding of how these scams operate and point to the ad networks as a potential place to intervene and demonetize this scam. We also mention that one of the ad networks identified in our study is currently being sued by the US Federal Trade Commission (FTC) for deceptive advertising practices. [3]

2 Background

At the core of the OSN spam ecosystem are the CPA affiliate-based ad networks that handle the task of monetizing the visitors generated by these spam-based abu-

[1] In these types of ad networks an advertiser only pays for the ad when the desired action has occurred. This action can range from the visitor installing a paid browser toolbar, purchasing some product, or providing their contact and personal information.

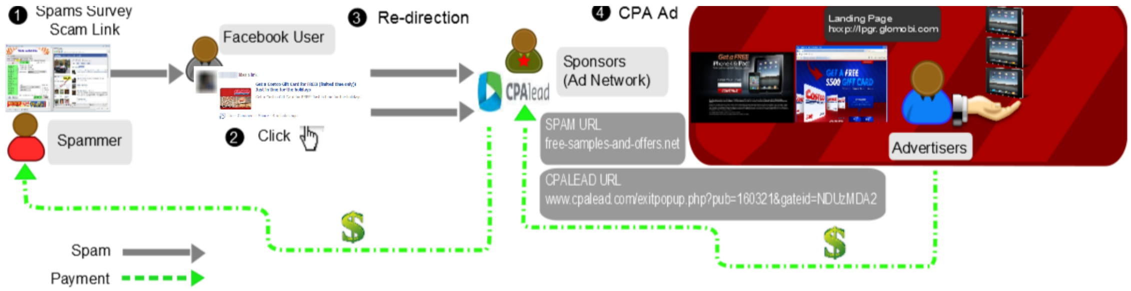
1 Spams Survey Scam Link
Facebook User
2 Click
3 Re-direction
4 CPA Ad
CPAlead
Sponsors (Ad Network)
Landing Page hxxp://lpgr.glomobi.com
Advertisers
SPAM URL free-samples-and-offers.net
CPALEAD URL www.cpalead.com/exitpopup.php?pub=160321&gateid=NDUzMDA2
Spammer
$
$
Spam
Payment

Figure 1: Showing the flow of money between the different entities.

sive advertising channels. These ad networks' presence frees the spammer from needing to deal with monetizing their victims and allows the spammers the ability to specialize in generating more effective spam campaigns. As detailed in previous research, the affiliate program is an efficient organizational model that decreases the risk to both parties and allows for greater flexibility and innovation [4]. As with most ad networks, these CPA ad networks are simply middle men in this scheme that line up advertisers that are responsible for creating and hosting the actual "surveys". These advertisers pay the ad network for each successfully performed action, and finally the ad network pays a fraction of this revenue to the affiliates that originally attracted the user to their ad network. Figure 1 depicts the typical chain of events and flow of money involved in these survey scams.

3 Data Sets and Methodology

Our exploration of the connection of ad networks and their sponsoring of affiliates that engage in Facebook spamming is based on two sources of data. The primary source of data for our investigation is a feed of Facebook spam detected by the MyPageKeeper Facebook application [6] and crawling data. Our secondary source is data from infiltrating ad networks that were identified as sponsoring abusive affiliates.

3.1 MyPageKeeper spam feed

From December 2012 until April 2013, we successfully crawled 1708 Facebook spam URLs that were identified by the MyPageKeeper Facebook application, which 2.2 million Facebook users located around the world have installed to protect their Facebook profiles from spam [6]. Of these, 17% (283) were survey scams, 50% (862) were broken links, 27% (458) were false positives, and 6% (105) were non-survey scams. This means that 73% (283 of the 388 working spam URLs) were survey scams. An inherent limitation of our study is the fact that our spam feed might contain a bias based on the algorithm used to detect spam posting and the user base that has installed MyPageKeeper. However, previous studies on the quality of spam feeds shows that user-based spam feeds as opposed to spam trap-based feeds tend to provide good coverage [5].

We then crawled these URLs using a webcrawler that is capable of following HTTP and Java-script redirection chains [1]. This crawling produced images of the page that we visually inspected to identify CPA ad network offers and that we manually interacted with to identify which ad networks were sponsoring each URL.[2] Based on crawling and manually collecting data we ended up with 1708 total URLs, 129 analyzed URLs, 93 unique landing domains, 32 observed Ad Networks, and 77 unique publisher IDs.

3.2 Infiltration

We attempted to register as an affiliate at each ad network that we encountered in order to gain additional insight into the ad networks.[3] Table 1 shows the names of the ad networks we encountered and whether or not we were able to register successfully.[4] It should be mentioned that there are at least three unique ways to measure/calculate prevalence. The first method is to count all of the raw URLs that we encountered. The second method is to count all of the unique landing pages and the third method is to count all of the publisher IDs. Each method has its own bias and drawback.

By infiltrating these ad networks, we were able to obtain two important pieces of ground truth information:

[2] Note that many of the initial spam URLs are URL-shortening services and lead to a smaller number of unique landing pages.

[3] Often, the ad networks required that we provide an explanation of our marketing methods, including our plans for driving traffic to the affiliate and a required upfront "interview."

[4] Sometimes language barriers were the primary reason we were unsuccessful at infiltrating the ad network, as was the case with ClickBanner, which is based in Greece. Other times the ad network was essentially closed and by invitation only.

URL	*Init* : *Land*	*URL*	*Init* : *Land*
007CPA	2 : 2	Fileice	2 : 2
A4D	4 : 3	Forestview	2 : 2
Ad.fly∗	3 : 3	Gurumedia	2 : 2
Adjal∗	1 : 1	LifeStreet Media∗	6 : 6
AdscendMedia∗	2 : 2	Lyris	1 : 1
AdvertMarketing	2 : 2	MaxBounty∗	1 : 1
Adworkmedia	2 : 2	Obey.my∗	2 : 1
Altervista∗	1 : 1	PulsePoint∗	2 : 2
Amung.us∗	35 : 33	Rapleaf∗	1 : 1
Aweber	1 : 1	ViralUrl∗	2 : 2
Bodis∗	1 : 1	W4	1 : 1
ClickBanner	13 : 4	Whitefire	1 : 1
Clicksor∗	6 : 1	YeahMobi∗	1 : 1
CPAlead∗	15 : 12	Zoosk	1 : 1
Escalatenetwork	1 : 1		:

Table 1: Summary of the prevalence of the affiliates calculated using the initial URL and Landing methods. ∗indicates that we successfully joined this ad network.

(1) links provided to their affiliates, which allowed us to extract affiliate and offer IDs as demonstrated below, and (2) understand how affiliate IDs are assigned, which we use in the results section to estimate the age of spammer's affiliate accounts.

3.3 Ad Network and Affiliate ID extraction

We manually interacted with these 129 unique landing pages and recorded network traffic traces. Analysis of these network traffic traces allowed us to identify the sponsoring ad network in most cases and we were able to identify the publisher or affiliate ID that belongs to an account that the spammer registered with the ad network.

In order for an affiliate to get credited with a completed survey, the ad network provides their affiliates with a URL that in most cases includes the affiliate's ID number and the offer ID. Table 2 provides two examples of the initial URL, the parsed affiliate and offer id, sponsor name, and the full ad network URL. Furthermore, for the programs we were able to infiltrate, we verified that our methods of identifying the ad network and extracting the affiliate and offer IDs were correct.

4 Results

In this section, we present some of our initial results on prevalence of ad networks, carbon dating, and revenue generation as it relates to survey scams.

4.1 Prevalence

Table 1 includes two metrics, unique spam URLs and unique landing pages, to estimate the prevalence of ad networks. Using either of these metrics, Amung.us ranks first and CPAlead is second. Both of our methods have different limitations, such as link shortener URLs are over counted in the case of unique spam URLs and landing pages are under counted if the landing page is the offer page instead of an intermediate page. Given these limitations, using the first metric of 129 unique spam URLs shows that over 50% of these URLs were traced back to four ad networks: Amung.us, CPAlead, ClickBanner, and LifeStreet Media.

4.2 Carbon Dating

In a previous study by Kanich et al. [2], they were able to estimate the revenue generated by illicit pharmacy affiliate programs using the insight that order IDs were sequentially allocated for each new order. We make use of a similar insight that affiliate IDs appear to be allocated sequentially in five of the ad networks identified. If our sequential affiliate ID allocation hypothesis is correct we use it along with some minimal ground truth data to "carbon date" (estimate the age of) affiliate IDs we extracted from the spam URLs.[5]

Via ad network infiltrations, we were able to obtain ground truth data of the age of two or more affiliate IDs spread out over time for eight ad networks. We then used this information to estimate the rate of affiliate ID allocation based on the increase in affiliate IDs we were allocated and the time that elapsed between registrations. If we assume that the rate of affiliate account creations is somewhat stable in the past, we can use the measured account registration rate to carbon date the affiliate IDs we have extracted.

We use the MaxBounty ad network as a concrete example of how our carbon dating methods works. When we initially infiltrated this ad network on 1-30-2013 we were assigned an ID of 123929 and when we joined MaxBounty for the second time on 4-22-2013 we received 129103 as our ID. To calculate the rate of affiliates joining the program, we take 82 days which is the time that elapsed between our first and second time joining MaxBounty, and the difference in ID assigned, which is 5174, and this results in a rate of 63/day affiliates joining MaxBounty. We observed MaxBounty affiliate ID 117373 in the scam feed which is 6556 less than our ID issued on 1-30-2013. Thus, the estimate from our carbon dating method is that the 117373 was issued on approximately 10-18-2012. To reinforce our results, we joined MaxBounty for a third time on 6-21-2013 and received 132424 as our ID. Therefore, we take 142 days which is the time elapsed between our first and third time. The 132424 ID is 8495 less than the initial

[5] In the case of CPAlead we confirmed by rapidly creating two affiliate accounts that we were allocated sequential affiliate IDs. Additionally, we have never observed an affiliate ID in the wild that was greater than one we were allocated within the corresponding time periods.

Initial URL	**Aff. ID**	**Offer ID**	**Ad Network**	**Ad Network URL**
`claimafreeiphone5.tk`	117373	5055	Maxbounty	`mb01.com/lnk.asp?o=5055&c=918273&a=117373`
`bit.ly/TygM3T`	1292	2199	007CPA	`track.007cpa.com/aff_c?offer_id=2199&aff_id=1292`

Table 2: Extraction of Affiliate ID, Offer Id, and Sponsor from Wireshark Capture of URL

Ad Network	Minimum	Maximum	Median	Average
Adscend	$0.11	$11.90	$0.63	$1.33
CPAlead	$0.03	$34.00	$1.05	$3.52
MaxBounty	$0.60	$3.75	$2.50	$2.33

Table 3: Offer Payouts for June 2013

ID of 123929 which results in a rate of 60/day affiliates joining MaxBounty. This rate would still put the carbon dating of the ID 117373 we observed in the scam feed to be within the week of 10-18-2012.

We acknowledge that our carbon dating method can only provide an estimate of an affiliate account's age. However, given this limitation we find that for the eight ad networks that we can compute their rate of affiliate account registrations the average spammer affiliate account age was approximately nine months old upon first observing this affiliate ID. There are two possible reasons that the majority of the spammer's affiliate accounts are old: 1) Spammers age their account before using them to avoid suspicion from the ad networks that are generating survey completions by spamming. 2) The ad networks are not doing a good job of detecting misbehaving affiliates that are engaging in abusive spamming activity.

4.3 Revenue Estimation

Ad networks allow affiliates to select which offer to direct users to from a wide number of offers and each one of these offers has a payout that is the amount of money paid to the affiliate for each successfully completed offer. Additionally, some ad networks provide the average conversion rate and Expected Payout per Click (EPC). For example, CPAlead offered a survey with the name "Airline Survey" with a payout of $1.24, an EPC of $0.02, and a conversion rate of 2%. Table 3 shows the minimum, maximum, average, and median payouts of survey offers from three ad networks.

5 Conclusions and Future Work

We will attempt to estimate how much traffic is generated via Facebook spam, some initial ideas of how to perform this include making use of the Bit.ly API that allows anyone to view how many times a shortened URL has been clicked. Other methods might include passive DNS measurements to estimate how many times a spammer's domain has been resolved. The key to combating these scams is to intervene at the ad syndication network level. This intervention might be self regulation, FTC and other international Government regulation, economic pressure from the advertisers, or some combination of the aforementioned techniques.

We presented an empirical study of Facebook spam revealing that 73% of the working spam URLs in our Facebook spam feed were monetized via survey scams sponsored by ad networks. Based on our analysis of 129 unique spam URLs over 50% of these URLs were traced back to four ad networks: Amung.us, CPAlead, ClickBanner, and LifeStreet Media. We presented a carbon dating method that can estimate the age of a spammer's affiliate ID and showed they are on average nine months old. Our preliminary results provide a potential point to demonetize the spam ecosystem by intervening on these deceptive ad networks that are sponsoring the majority of Facebook spam seen in our study.

Acknowledgments

We thank Harsha Madhyastha and Sazzadur Rahman for their help with this project. This work was supported by the National Science Foundation under grant 1237076. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsors.

References

[1] KANICH, C., CHACHRA, N., ET AL. No plan survives contact: Experience with cybercrime measurement. *Proc. of 4th USENIX CSET* (2011).

[2] KANICH, C., WEAVER, N., ET AL. Show Me the Money: Characterizing Spam-advertised Revenue. In *Proceedings of the USENIX Security Symposium* (San Francisco, CA, August 2011).

[3] LATTIN, P. FTC Pounds Hard Rishab Verma of OBEY. http://performinsider.com/2013/03/ftc-pounds-hard-rishab-verma-of-obey/, 2013.

[4] LEVCHENKO, K., PITSILLIDIS, A., ET AL. Click trajectories: End-to-end analysis of the spam value chain. In *IEEE Symposium on Security and Privacy* (Washington, DC, USA, 2011), SP '11, IEEE Computer Society.

[5] PITSILLIDIS, A., KANICH, C., ET AL. Taster's choice: a comparative analysis of spam feeds. In *IMC '12* (2012).

[6] RAHMAN, M., HUANG, T., MADHYASTHA, H., AND FALOUTSOS, M. Frappe: detecting malicious facebook applications. In *Proceedings of the 8th international conference on Emerging networking experiments and technologies* (2012), ACM, pp. 313–324.

[7] THOMAS, K., AND NICOL, D. M. The koobface botnet and the rise of social malware. In *MALWARE 2010* (2010).