May 21st, 11:30 AM

Tracking Criminals on Facebook: A Case Study From A Digital Forensics REU Program

Daniel Weiss
University of Arizona, dweiss@email.arizona.edu

Gary Warner
University of Alabama at Birmingham, Director of Research in Computer Forensics

Follow this and additional works at: https://commons.erau.edu/adfsl

Part of the Aviation Safety and Security Commons, Computer Law Commons, Defense and Security Studies Commons, Forensic Science and Technology Commons, Information Security Commons, National Security Law Commons, OS and Networks Commons, Other Computer Sciences Commons, and the Social Control, Law, Crime, and Deviance Commons

Scholarly Commons Citation
https://commons.erau.edu/adfsl/2015/thursday/4

This Peer Reviewed Paper is brought to you for free and open access by the Conferences at Scholarly Commons. It has been accepted for inclusion in Annual ADFSL Conference on Digital Forensics, Security and Law by an authorized administrator of Scholarly Commons. For more information, please contact commons@erau.edu.
ABSTRACT

The 2014 Digital Forensics Research Experience for Undergraduates (REU) Program at the University of Alabama at Birmingham (UAB) focused its summer efforts on tracking criminal forums and Facebook groups. The UAB-REU Facebook team was provided with a list of about 60 known criminal groups on Facebook, with a goal to track illegal information posted in these groups and ultimately store the information in a searchable database for use by digital forensic analysts. Over the course of about eight weeks, the UAB-REU Facebook team created a database with over 400 Facebook groups conducting criminal activity along with over 100,000 unique users within these groups. As of November 2014, students involved in the research project with Advisor Gary Warner at UAB continued running the automated fetchers since my summer project completed. Working with U.S. Federal Law Enforcement agencies, there have been at least NINE CONFIRMED ARRESTS of individuals associated with the illegal activities tracked on Facebook. This paper will discuss the methods used to collect the information, store it in a database and analyze the data. The paper will also present possible future uses of the Facebook criminal activity-monitoring tool.

Keywords: social media, criminal organizations, online crime, social network monitoring

1. INTRODUCTION

For the past five years, the UAB Computer Forensics Research Lab has participated in the National Science Foundation Research Experience for Undergrads (REU) program. During the summer of 2014, the Digital Forensics REU team focused on developing tools for automating the gathering and analysis of the communications between criminals in online forums and on Facebook groups. The UAB-REU summer 2014 research project created a searchable database that keeps track of the growing criminal activity on Facebook. Our case study has a growing database that can keep track of everything on a Facebook group from posts, comments, likes, as well as the user who posted the respective post, the time it was posted, and any image that was posted in a post or comment. This data can be used to draw connections between active users within different groups and lead to arrests if proven criminal acts were performed. Many of the messages that we stored within the database contained credit card numbers associated with other personal information as well.

2. LITERATURE REVIEW

Previous REU cohorts have examined the methods in which criminals learn and encourage one another’s criminal behavior through online social interaction in the area of phishing. (Levin, Richardson, Warner, & Kerley, 2012) Others have explored the role of online social media networks in the creation and execution of large international markets for stolen data and identities. Several researchers have examined online web forums that were designed primarily to support international trade in stolen goods and identities. (Holt & Smirnova, 2014), (Motoyama, McCoy,
Levchenko, Savage, & Voelker, 2011), (Merces, 2011) As criminals and terrorist grow more brazen, they have realized that the use of secretive online forums is not necessary when Facebook traffic is largely unregulated and unmoderated and represents minimal risk of prosecution or incarceration. The House Homeland Security Committee held hearings on “Jihadist Use of Social Media” in 2011 where testimony included “The Antisocial Network” where it was remarked how little concern adversaries have about discovery. (Kohlman, 2011)

The Law Reviews and Journals are beginning to fill with articles about the use of evidence from social media in the courts. Many of the opinions expressed in those articles helped to make the case for the existence of this project. One current trend in this debate is whether messages shared “quasi-privately” only to a chosen community of friend’s withstood Fourth Amendment challenges regarding expectations of privacy. (Sholl, 2013) Others have argued about the admissibility of such evidence, partly with regards to whether it constituted heresay under Federal Rules of Evidence. (Holt & San Pedro, 2014) Still others argue about the authentication of the evidence and how to prove the origins and identify of the poster. (Griffith, Winter 2012).

To address all of these concerns, evidence would need to be gathered in a repeatable and automated way that preserved the timestamp and ‘userid’ of the creator of the evidence, and only from pages that could be shown to be publicly “Open.”

3. FACEBOOK AS OPEN SOURCE INTELLIGENCE

3.1. Problem Statement Summary

The UAB-REU Facebook team was given a list of known criminal groups on Facebook, and was asked to track these groups over the summer of 2014. Specifically, the following was to be accomplished by the end of the summer. Can we quickly decide if a Facebook group is discussing criminal activity and if so, can we characterize what types of activities they do or targets they are after. For example is the criminal activity credit card fraud, stolen electronics, shipping of illegal or stolen items, viruses, malware, botnets, spamming, and even terrorists organizations or supporters of terrorists. We also wanted to be able to identify the most influential, and or important people, and or most active users within a group. By the end of summer our goal was to be tracking at least 200 criminal Facebook groups. With these goals in mind we set out to develop code to request and retrieve the wanted information from Facebook, and store the information into a searchable database where we could easily query the data for further investigations.

3.2. Facebook’s Graph Application Programming Interface (API)

The API is on the developer side of Facebook and is a great tool that we used over the summer project. “The graph API is the primary way to get data in and out of Facebook’s social graph (network).” Essentially the Graph API allows a user to post, delete, and also get information to and from Facebook. The graph API was a tremendous asset for our team because it allowed us to query many useful searches directly without having to perform many iterations to gather wanted information, however to do so an Access Token was required.

3.2.1. The Basics

The Graph API is a representation of the information on Facebook, which is composed of nodes, edges, and fields. Nodes are basically the “things” on Facebook. Ex. Users, Photos, Posts. Edges are the connections between nodes, such as a comment or a like on a photo. Fields are the information about nodes. For example, a node that is a user can have a field such as their birthday or hometown.

3.2.2. Using the Graph API to find more criminal groups

To find more criminal Facebook groups, we used the Graph API, and searched for groups with specific keywords. Group names that had the word such as “Hacker” or “CVV” within their

1 Graph API Overview
https://developers.facebook.com/docs/graph-api/overview
name were added to our list of criminal groups. Even though it was not for sure that these groups were criminal our database queries later on would tell us. Figure 1 below shows the Graph API searching for all groups with the word “Hacking” in its’ name. Our team developed a “Bag of Words” which essentially was a list of keywords that we used to find new Facebook groups.

![Graph API Search](https://developers.facebook.com)

**Figure 1 Graph API Search**

Source: [https://developers.facebook.com](https://developers.facebook.com)

### 3.2.3. Facebook Privacy

The Graph API is a very handy tool that Facebook has allowed the public to use. However, Facebook privacy still comes into play when using the API. Facebook groups that have a Privacy status of Open, meaning anyone can see the group and join it, or a status of Closed, meaning anyone can see the group but must request to join the group can be seen through the Graph API. A group that is secret will not show up on the API. A secret group has no record of existing through any means of searches; the only way to be in a secret group is by getting invited to join the group. Of course, being in a closed or secret group allows users to see everything going on within the group making the group ‘Open’ to the users within. Figure 2 and Figure 3 below are examples of an open and a closed group. Notice the difference in the amount of information between the Open and Closed group. Figure 4 below is an example of a closed group that the current Facebook user on the Graph API was a member of. Notice that it now looks like an open group.
Figure 2 Open Group Example
Source: https://developers.facebook.com

Figure 3 Closed Group non-Member Example
Source: https://developers.facebook.com
3.2.4. Aliases
When collecting group ID numbers to run through the fetcher we realized that we were only able to pull information from a group that was open. To fix this issue we created Facebook aliases that looked like cyber criminals. We made two main accounts in particular and tried joining as many of the closed groups that we had found through the Graph API as possible. As a matter of fact it was not very hard to get accepted to a number of these groups. Once accepted into these groups we would run the Graph API with our alias’ Access Token and then run the fetcher. This was a huge step in our summer research as it allowed us to gather a considerably larger amount of data.

4. CODE IMPLEMENTATION

4.1. Automation of the Graph API
The program for extracting the information from Facebook was written in Java. The code used a library package called RestFB, which allowed for direct access to the Graph API while in Java. We would supply the Graph API with a Group ID number and then would retrieve all of the group’s members, posts, comments, likes, pictures, etc.

4.2. The Database
Our team created an SQL database to store the data retrieved from Facebook and make it easy to search for wanted results. In SQL, several different tables were created to easily make connections between users and groups. Within Java we coded to put all of the comments within its own SQL searchable table for example. Similarly tables were used to store information for images, posts, and groups. We created a user group’s table that allowed us to connect users to multiple groups because there were many instances where the same user belonged to more than one group in our database, and this allowed for a connection between the two.

5. RESULTS
Within the database we ran queries to achieve the goals we set out in the beginning of the summer. We were able to determine if a Facebook group was talking about criminal activity, what kind of activity, and the ‘big’ players within those groups.
as well. By the end of the summer after about two
weeks of data collection, the database had over
400 criminal groups that we were tracking and
fetching information from. Within those 400
groups, there were over 100,000 unique users in
those groups, about 50,000 posts, and about
40,000 comments on posts.

The following query looked for messages within
the group’s posts that contained a certain
keyword. The query searched for posts containing
the word ‘fbi’. Many other related queries
searched for posts containing the words ‘cia’ or
‘vbv’ (Verified By Visa, a common term used by
credit card criminals.) i.e. and counted the number
of occurrences, displaying then the top ten groups.

<table>
<thead>
<tr>
<th>count</th>
<th>Groupid</th>
<th>name</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>183381435133188</td>
<td>SPAMER’s</td>
</tr>
<tr>
<td>11</td>
<td>229655640389234</td>
<td>KING OF HACKER</td>
</tr>
<tr>
<td>10</td>
<td>505516012807000</td>
<td>DDOS</td>
</tr>
<tr>
<td>9</td>
<td>230749693747529</td>
<td>! P4K OR4Kz4I H4CkERX !</td>
</tr>
<tr>
<td>8</td>
<td>465238643517306</td>
<td>Bestbikes Grupo ventas Nacional</td>
</tr>
<tr>
<td>8</td>
<td>165155633573484</td>
<td>WESTERN UNION</td>
</tr>
<tr>
<td>7</td>
<td>290630927627110</td>
<td>Genius Hackers</td>
</tr>
<tr>
<td>6</td>
<td>126115430891994</td>
<td>SaDaM Khakwani All Hacking TrickXx &amp; Tip$</td>
</tr>
<tr>
<td>6</td>
<td>112852328867059</td>
<td>HACKERS SPOTTED :)</td>
</tr>
<tr>
<td>6</td>
<td>14929934514034</td>
<td>Hack With Stylee (Hacking Zone)</td>
</tr>
</tbody>
</table>

The following query searched for messages that
contained a string of 15 or 16 digits, because that
was our credit card number identifier if groups
were sharing stolen credit cards with one another.
The query below shows the results for the top four
group’s sharing Visa credit cards. Our query
searched for the number four followed by another
15 digits 0-9.

<table>
<thead>
<tr>
<th>count</th>
<th>Fb_group.Groupid</th>
<th>Fb_group.name</th>
</tr>
</thead>
<tbody>
<tr>
<td>432</td>
<td>435715723187958</td>
<td>PRO SHOPPER’S TUT AND BINS AND STORES</td>
</tr>
<tr>
<td>402</td>
<td>563652277096630</td>
<td>REEF GH <em><strong>CCV STRONG CARDS</strong></em> KILL THEM ALL</td>
</tr>
<tr>
<td>376</td>
<td>384945978297975</td>
<td>PRO SHOPPERS ***KILL, WAL,KMAR,SEAR, AND BEST</td>
</tr>
</tbody>
</table>
The following query took a group that talked about visa credit card numbers frequently and displayed the message along with the user who posted it. (Card numbers have been altered for privacy.)

### Table 3: Results for the Visa query

<table>
<thead>
<tr>
<th>userid</th>
<th>name</th>
<th>substring</th>
</tr>
</thead>
<tbody>
<tr>
<td>100008366380917</td>
<td>Nana Less</td>
<td>4266841341509999 02/17 597 Sue Lowe 123 sixth street Calvin LA 71410</td>
</tr>
<tr>
<td>100008366380917</td>
<td>Nana Less</td>
<td>418586411539999 06/16 417 Debra Duhon 300 Big Pasture Rd Lake Charles LA</td>
</tr>
<tr>
<td>10000835312440</td>
<td>Okoeokoso More-vimlated Vim-carders</td>
<td>high balance cc 4347696620159999 1016 919 Cynthia Kroeker 11817 SW 1st Yukon OK 73099</td>
</tr>
<tr>
<td>100005869085570</td>
<td>Undergrad Carder</td>
<td>428208712259999 1014 578 Martin Ibarra 1108 E ORTEGA ST Santa Barbara C</td>
</tr>
</tbody>
</table>

### 6. Future Uses

After just a short eight weeks in this REU program, and after only two weeks of actual data collection, results were huge. As of November 2014, students involved in the research project with Advisor Gary Warner at UAB continued running the automated fetchers since my summer projected completed. After the REU program completed for the summer, the tool became the anchor of a new Open Source Intelligence effort within the lab. The database now contains over a half million Facebook messages and replies and is monitoring more than 900 Facebook groups. The most prolific of these groups that were found to be dedicated to criminal activity have logged well over 5,000 messages each from as many as 1,800 distinct Facebook users. The tool has been used to learn more about criminal groups for many Federal, state, and local law enforcement investigations. Original conceived to assist in cybercrime cases, investigations have included tracking of many types of Facebook groups including “carders” (criminals who steal and trade credit cards), “booters” (criminals who sell DDOSing services), online sexual harassment via webcam-controlling botnets, street gangs selling illegal drugs and weapons, and counter-terrorism investigations. Hundreds of Facebook groups have been reported and terminated, while others are left intact to identify ring-leaders and, working with major US-based shipping companies and retailers, to intercept the shipment of stolen packages. Working with an inter-agency task force on violent crime, Facebook evidence from this project was used to document relationships between criminals as well as proof of weapons and drug possession from photos shared on Facebook in support of a RICO case that led to nine felony arrests.

The project has also led to additional publications that have been focused on image analysis of the profile pictures. Hackers often use Guy Fawkes masks in profiles pictures, carders often have images of credit cards in their profile pictures, and jihadists often have Islamic State flags on their profile pictures. In addition to keyword clues,
these new image analysis tools allow a group to be quickly categorized, even when the language used in the messages is not understood by the analyst. Implementation of a tool like this would have a great impact on the cyber world, as it would aid in the capture of cyber criminals.

7. CONCLUSION

The 2014 Digital Forensics REU program at UAB provided students with the opportunity to develop real world applications with valuable outcomes. Our 2014 project identified criminal activity on Facebook, collected evidence and ultimately helped prosecute and punish criminals. The UAB-REU Facebook team created a searchable database that could be used by law enforcement and intelligence agencies, as well as private sector shipping companies, banks, and credit card companies to identify criminal activity and work with law enforcement to prosecute those responsible for the illegal activity.

REFERENCES

1. BIBLIOGRAPHY


accounts. *Proceedings of the second symposium on Usable privacy and security* (pp. 44-45). ACM.


Environmental, and International Guidance and Best Practices, 267.


