Efficient and Scalable Socware Detection in Online Social Networks

Md Sazzadur Rahman, Ting-Kai Huang, Harsha Madhyastha, Michalis Faloutsos

University of California, Riverside
Problem Statement

- Social malware is rampant on Facebook

Cyber Attacks Jump 81 Percent, Target Mobile, Social Networks

83 million Facebook accounts are fakes and dupes

Charlie Sheen death hoax spreads malware through Facebook

Facebook Scam: “Southwest Free Flights” Will Comment Spam Your Friends
Socware – a new kind of malware

• Social Malware = Socware
• New propagation platform
  – Leverage users for viral propagation unlike email spam
Socware – needs new defenses

• Often hosted inside Facebook domain
  – No blacklist designed to identify malicious apps

Avengers

ABOUT THIS APP
Who can see posts this app makes for you on your Facebook timeline: [?]  
Friends of Friends [ ]

THIS APP WILL RECEIVE:
• Your basic info [?]

Report App

This app may post on your behalf, including status updates, photos and more.
Contributions

• Develop MyPageKeeper, a security app in Facebook
  – Obtained 12K users, effectively monitors 2.4M wall
• Design efficient, scalable socware detection method
  – 0.005% false positive, 0.3% false negative rates
  – 12 EC2 micro instances monitors 12K users every 2 hours
• Socware is a new kind of malware
  – 26% of flagged posts point back to Facebook
Roadmap

• MyPageKeeper
• Evaluation
• Interesting results
• Conclusion
MyPageKeeper: Facebook security app

Launched July, 2011

https://apps.facebook.com/mypagekeeper/
MyPageKeeper operations

1. Obtaining user’s authorization
2. Crawling users wall and news feed
3. Identifying socware and notification
Classifying socware

• Previous approaches
  – URL crawling based approach [Thomas, IEEE S&P’11]
    • Expensive (large network latency)
  – Lexical, host properties of landing URL [Ma, SIGKDD’09]
    • Socware often hosted inside Facebook
Classifying software

• Our approach for fast and effective classification:
  – Features readily available in a post
  – Aggregate features across users

• Train Support Vector Machine classifier:
  – Spam keyword score
  – Message similarity
  – No of news feed and wall post
  – No of Like and comments
  – URL obfuscation
Spam keyword score

Awesome, Free, wow, OMG…
Message similarity

Similar text in socware posts

WOW I can't believe that you can see who is viewing your Profile! I just saw My Top 10 Profile and Photo Peekers and I am SHOCKED!! You can also see WHO VIEWED YOUR PROFILE here --> http://bit.ly/M9cf8y

WOW I can't believe that you can see who is viewing your Profile! I just saw My Top 10 Profile and Photo Peekers and I am SHOCKED!! You can also see WHO VIEWED YOUR PROFILE here --> http://bit.ly/M9cf8y

WOW I can't believe that you can see who is viewing your Profile! I just saw My Top 10 Profile and Photo Peekers and I am SHOCKED!! You can also see WHO VIEWED YOUR PROFILE here --> http://bit.ly/M9cf8y

Different text in non-socware posts

They sound very good I like them http://www.youtube.com/watch?v=d9NF2edx-M&feature=share


I know the economy is bad, but canada has government funded musical programs, i'me sure these guys could've bought more instruments. They do kill this performance though http://www.youtube.com/watch?v=d9NF2edx-M&feature=share

Check this version out. http://www.youtube.com/watch?v=d9NF2edx-M&feature=share
Other features

• News feed post and wall post count
• Like and comment count
• URL Obfuscation
Implementation

• Non-trivial engineering
  – Apache, Django, Postgres

• Designed for scale
  – Crawler instances
    • 12 micro instances in Amazon EC2 to run crawlers
  – Checker instances
    • Work load partitioned across checkers
Roadmap

- MyPageKeeper
- Evaluation
- Interesting results
- Conclusion
## Evaluation: Dataset

Collected from June to October 2011

<table>
<thead>
<tr>
<th>Data</th>
<th>Total count</th>
<th># of distinct URLs</th>
</tr>
</thead>
<tbody>
<tr>
<td>MyPageKeeper users</td>
<td>12.5K</td>
<td>-</td>
</tr>
<tr>
<td>Friends of MPK users</td>
<td>2.4M</td>
<td>-</td>
</tr>
<tr>
<td>News feed posts</td>
<td>38.8M</td>
<td>29.5M</td>
</tr>
<tr>
<td>Wall posts</td>
<td>1.8M</td>
<td>1.5M</td>
</tr>
<tr>
<td>User reports</td>
<td>679</td>
<td>333</td>
</tr>
</tbody>
</table>
Is MyPageKeeper accurate?

- **Training**
  - 2500 positive and 5000 negative sample posts

- **Testing**

  
  - 4,972 unique URLs in 60,191 posts
  - 752,516 unique URLs

  40M posts, 31.5M unique URLs

  Post DB
MyPageKeeper is accurate

97% software identified by classifier true positives.

Flagged by Blacklists
Flagged by on.fb.me
Content deleted by Facebook
Blacklisted app
Blacklisted IP
Domain is deleted
Points to app install
Spamming app
Manually verified
True positives
Unknown
Total

# of posts
MyPageKeeper is efficient

20x better throughput than other approaches
Blacklist does not suffice
96.6% malware identified by classifier

<table>
<thead>
<tr>
<th>Source</th>
<th>% of posts</th>
<th>Overlap with classifier (# of URLs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google SBA2</td>
<td>0.4%</td>
<td>0</td>
</tr>
<tr>
<td>Phishtank</td>
<td>0.5%</td>
<td>1</td>
</tr>
<tr>
<td>Malware Norm</td>
<td>0.2%</td>
<td>0</td>
</tr>
<tr>
<td>Joewein</td>
<td>0.7%</td>
<td>11</td>
</tr>
<tr>
<td>APWG</td>
<td>0.6%</td>
<td>0</td>
</tr>
<tr>
<td>Spamcop</td>
<td>1.0%</td>
<td>0</td>
</tr>
<tr>
<td><strong>All blacklists</strong></td>
<td><strong>3.4%</strong></td>
<td>12</td>
</tr>
<tr>
<td><strong>MyPageKeeper Classifier</strong></td>
<td><strong>96.6%</strong></td>
<td>-</td>
</tr>
</tbody>
</table>
Roadmap

• MyPageKeeper
• Evaluation
• Interesting results
• Conclusion
Socware is prevalent in Facebook

- 49% MyPageKeeper users exposed to socware
- 26% of socware post points back to Facebook
Novel Parasitic behavior: Like-as-a-Service

Artificially inflate number of Likes in Facebook page
Novel Parasitic behavior: Like-as-a-Service

- Identified 721 distinct LaaS URLs in Facebook
- Other Like-as-a-Service app: Games, FanOffer, Latest Promotion
Conclusion

• We presented MyPageKeeper, a security app
• We show that
  – Reach of socware is widespread
  – Existing defenses not sufficient to identify socware
  – MyPageKeeper is accurate and efficient
• We identified LaaS, new aggressive marketing trend
Thank you!

Questions?

https://apps.facebook.com/mypagekeeper/
http://mypagekeeper.org