Using Machine-Learning to Investigate Web Campaigns at Large

Dr. Federico Maggi, @phretor

Joint work with Marco Balduzzi, Ryan Flores, Lion Gu and Vincenzo Ciancaglini
Web Defacement = Website Compromise + Homepage Alteration
FOR IMMEDIATE RELEASE

Thursday, May 10, 2018

California Man Arrested For Hacking Websites For The Combating Terrorism Center At West Point And The New York City Comptroller

The Defendant Committed More Than 11,000 Defacements of Various Military, Government, and Business Websites Around the World Using the Online Pseudonym “Alfabetovirtual”

Geoffrey S. Berman, the United States Attorney for the Southern District of New York, and William F. Sweeney Jr., the Assistant Director-in-Charge of the New York Field Office of the Federal Bureau of Investigation (“FBI”), announced today the arrest of BILLY RIBEIRO ANDERSON, a/k/a “Anderson Albuquerque,” a/k/a “Alfabetovirtual.” ANDERSON was charged with three separate counts of computer fraud for obtaining unauthorized access to and committing defacements of the websites for the Combating Terrorism Center at the United States Military Academy in West Point, New York (“West Point”), and the Office of the New York City Comptroller (the “NYC Comptroller”). ANDERSON was arrested earlier this
Meerkat: Detecting Website Defacements through Image-based Object Recognition

Kevin Borgolte, Christopher Kruegel, Giovanni Vigna
University of California, Santa Barbara
{kevinbo,chris,vigna}@cs.ucsb.edu

Abstract
Website defacements and website vandalism can inflict significant harm on the website owner through the loss of sales, declines in reputation, and increased moderation costs. Meerkat, a novel approach to detecting defacement, builds upon an image-based object recognition model to identify alterations to a website.

1 Introduction
The defacement and vandalism of websites is an attack that disrupts the operation of companies and organizations, tarnishes their brand, and plagues websites of all sizes, from those with personal content to large corporation sites.
Evolution

“Just-for-fun” era

Attention to geo-political events

90s

2000

2010

Now

Digital activism

Dark propaganda
“Just-for-fun” Era

HACKED BY LIBERO

Login : Libero   Ok!
Password: **********   Ok!
Conection Ok!

CFK !!!

WE ARE ARGENTINA HACK TEAM

| owned by ssh-2 | CHILE |

uname -a;id

Linux obelix 2.6.15-1-686 #2
Mon Mar 6 15:27:08 UTC 2006
i686 GNU/Linux

uid=0(root) gid=33(www-data) groups=33(www-data)
Evolution

“Just-for-fun” era

Attention to geo-political events

90s

Digital activism

2000

2010

Now

Dark propaganda
We hack because we are fighting against bad information, all of us ain't terrorists, we're curious. Is it a crime to be interested in learning? Each of us has an own life, which may be good or bad. This above are places where freedom is a dream, open your eyes, open your mind, wake up, and all together we'll turn freedom from dream to reality. You can agree with us or not. If you agree with us let diffuse this message to make our ideas become everybody's ideas.
Evolution

“Just-for-fun” era

90s

2000

Digital activism

2010

Dark propaganda

Now

Attention to geo-political events
SANTIAGO, Chile — Gen. Augusto Pinochet, who terrorized his opponents for 17 years after taking power in a bloody coup, died Sunday, putting an end to a decade of intensifying efforts to bring him to trial for human rights abuses blamed on his regime. He was 91.

Supporters saw Pinochet as a Cold War hero for overthrowing democratically elected President Salvador Allende at a time when the U.S. was working to destabilize his Marxist government and keep Chile from exporting communism in Latin America.
Evolution

“Just-for-fun” era

Attention to geo-political events

90s

2000

2010

Now

Digital activism

Dark propaganda
From isolated events to coordinated campaigns

- Death statement is June 25th, 2009
- Sporadic defacements: June 28th
- Campaigns: August

Own3d by FIREH4CK3R - Fail Shell

"Se queremos um bom país vamos amá-lo mais!
Seja mais você e não os outros!
Tenha orgulho de ser brasileiro!"

We are: FIREH4CK3R - Hackinho - Crackinho - Twi John

LUTO MICHAEL JACKSON
That's interesting
But how do we scale?
# Public Repositories

## TOP30 User

<table>
<thead>
<tr>
<th>No</th>
<th>User</th>
<th>PR</th>
<th>View</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>越南邻国宰相 [34333]</td>
<td>0</td>
<td>View</td>
</tr>
<tr>
<td>2</td>
<td>Jack Riderr [19199]</td>
<td>0</td>
<td>View</td>
</tr>
<tr>
<td>3</td>
<td>Mr.Kro0oz.305 [10672]</td>
<td>0</td>
<td>View</td>
</tr>
<tr>
<td>4</td>
<td>AlfabetoVirtual [8907]</td>
<td>0</td>
<td>View</td>
</tr>
<tr>
<td>5</td>
<td>宇少 [6559]</td>
<td>0</td>
<td>View</td>
</tr>
<tr>
<td>6</td>
<td>大圣 [6486]</td>
<td>0</td>
<td>View</td>
</tr>
<tr>
<td>7</td>
<td>丿血 [5408]</td>
<td>0</td>
<td>View</td>
</tr>
<tr>
<td>8</td>
<td>That is me [5282]</td>
<td>0</td>
<td>View</td>
</tr>
<tr>
<td>9</td>
<td>悔悴 [5013]</td>
<td>0</td>
<td>View</td>
</tr>
<tr>
<td>10</td>
<td>頹废 [4791]</td>
<td>0</td>
<td>View</td>
</tr>
</tbody>
</table>

## Date | Notifier | Domain |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2016-02-12</td>
<td>Tr</td>
<td><a href="http://bananabags.co.za/">http://bananabags.co.za/</a></td>
</tr>
<tr>
<td>2016-02-12</td>
<td>Tr</td>
<td><a href="http://bsvreault.co.za/">http://bsvreault.co.za/</a></td>
</tr>
<tr>
<td>2016-02-12</td>
<td>Tr</td>
<td><a href="http://bsgtriathlon.org/">http://bsgtriathlon.org/</a></td>
</tr>
<tr>
<td>2016-02-12</td>
<td>Tr</td>
<td><a href="http://broadbandondemand.co.za/">http://broadbandondemand.co.za/</a></td>
</tr>
<tr>
<td>2016-02-12</td>
<td>Tr</td>
<td><a href="http://bettacars.co.za/">http://bettacars.co.za/</a></td>
</tr>
<tr>
<td>2016-02-12</td>
<td>Tr</td>
<td><a href="http://blog.camelotkennels.co.za/">http://blog.camelotkennels.co.za/</a></td>
</tr>
<tr>
<td>2016-02-12</td>
<td>Tr</td>
<td><a href="http://www.conseils-incendie.fr/images/j">http://www.conseils-incendie.fr/images/j</a></td>
</tr>
<tr>
<td>2016-02-12</td>
<td>Tr</td>
<td><a href="http://portalmaria.it/images/jdownloads/s">http://portalmaria.it/images/jdownloads/s</a></td>
</tr>
<tr>
<td>2016-02-12</td>
<td>Tr</td>
<td><a href="http://www.fruitinfo.hu/images/jdownload">http://www.fruitinfo.hu/images/jdownload</a></td>
</tr>
<tr>
<td>2016-02-12</td>
<td>Tr</td>
<td><a href="http://578h%E9%90%A8%E5%8B%86%E6%B6%93%E2%82%AC">http://578h鐨勆涓€</a></td>
</tr>
</tbody>
</table>
Hi There!

I am so happy to announce the new Mirror-zone.org unrestricted information site, with few nice changes.

We are now at testing this script, this means the script is not stable, and you may get some errors. If you face any issues please contact us and we will be glad to work with you to in order to resolve such issues. Also, this not mean that the Gold release will have the some futures with the bugs fixes, and we will continue to add new futures to the site.

Best Regards,
Administration

Hello

mirror-zone.org followers,
Did your defaced site mirror goes to onhold ?
Because your notify name are missing on your deface page

-------------------

Welcome to
MyDeface.com || Let's Show off....

Latest Verified Mirrors

<table>
<thead>
<tr>
<th>Name</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>lemon</td>
<td><a href="http://upsanddowns-horde">http://upsanddowns-horde</a></td>
</tr>
<tr>
<td>lemon</td>
<td><a href="http://bradfordourtainsandb">http://bradfordourtainsandb</a></td>
</tr>
<tr>
<td>lemon</td>
<td><a href="http://leatherheaddistrictscc">http://leatherheaddistrictscc</a></td>
</tr>
<tr>
<td>lemon</td>
<td><a href="http://leatherheaddistrictscc">http://leatherheaddistrictscc</a></td>
</tr>
<tr>
<td>lemon</td>
<td><a href="http://pickwickassociation">http://pickwickassociation</a></td>
</tr>
<tr>
<td>lemon</td>
<td><a href="http://rhodapartridgec">http://rhodapartridgec</a></td>
</tr>
</tbody>
</table>
THE site :)
## Collected Data

<table>
<thead>
<tr>
<th>Source Name</th>
<th>Website URL</th>
<th>Acquired Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zone-H</td>
<td><a href="http://www.zone-h.org">www.zone-h.org</a></td>
<td>12,303,240</td>
</tr>
<tr>
<td>Hack-CN</td>
<td><a href="http://www.hack-cn.com">www.hack-cn.com</a></td>
<td>386,705</td>
</tr>
<tr>
<td>Mirror Zone</td>
<td><a href="http://www.mirror-zone.org">www.mirror-zone.org</a></td>
<td>195,398</td>
</tr>
<tr>
<td>Hack Mirror</td>
<td><a href="http://www.hack-mirror.com">www.hack-mirror.com</a></td>
<td>68,980</td>
</tr>
<tr>
<td>MyDeface</td>
<td><a href="http://www.mydeface.com">www.mydeface.com</a></td>
<td>37,843</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td><strong>12,992,166</strong></td>
</tr>
</tbody>
</table>
Amount of Events Over Time

![Bar chart showing the amount of events over time from 1998 to 2016. The y-axis represents the number of defacements, ranging from 0 to 1,800,000. The x-axis represents the years from 1998 to 2016. The chart shows a significant increase in events after 2011, peaking in 2013 and 2014.]
Data Format

Metadata

Raw content

```
[ Xai Syndicate - ()

~~ Mr.DreamX196 - D4RKNE55 - 0xd3vs - ML7C - ./51N1CH1 - ./R015 - Mr.AchanX48 - Mr.BucketHead - Angel Dot Id - Ups.137 --
[< Littlebear69 - L3mot_n3t - Laser69 - Gend3ruw0 - SPEEDY-83 - megelang9etar - civiliant - KATENBAD - ./RosesDie - indonesia6etar! - Mr.Lucifers - Con7ext - LCR999X - Mr.7z>]
```
## Data Accuracy

<table>
<thead>
<tr>
<th>Type</th>
<th>Attribute</th>
<th>Example</th>
<th>Trustworthiness</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Metadata</strong></td>
<td>URL</td>
<td><a href="http://target.gov">http://target.gov</a></td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>Timestamp</td>
<td>2010-01-02 15:00</td>
<td>Medium</td>
</tr>
<tr>
<td></td>
<td>Nickname</td>
<td>Neo Hacker</td>
<td>Medium-Low</td>
</tr>
<tr>
<td></td>
<td>Webserver; Reason; Hack Mode</td>
<td>Nginx; Political; SQLi</td>
<td>Low</td>
</tr>
<tr>
<td><strong>Raw content</strong></td>
<td>Main page</td>
<td>HTML or TXT file</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>Embeeeded resources</td>
<td>Various format</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>External resources</td>
<td>Various format</td>
<td>Medium-High</td>
</tr>
</tbody>
</table>
General Trends
<table>
<thead>
<tr>
<th>Year</th>
<th>Most relevant topics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>question, student, security, number, place</td>
</tr>
<tr>
<td>1999</td>
<td>cowboy, team, security, think</td>
</tr>
<tr>
<td>2000</td>
<td>baby, tabloid, people, provided</td>
</tr>
<tr>
<td>2001</td>
<td>lord, prime, provided, saved, better</td>
</tr>
<tr>
<td>2002</td>
<td>worry, sind, lame, care, encryption</td>
</tr>
<tr>
<td>2003</td>
<td>backup, gift, team, came, take</td>
</tr>
<tr>
<td>2004</td>
<td>best, group, micro, look, total</td>
</tr>
<tr>
<td>2005</td>
<td>normal, pope, time, familia, contact</td>
</tr>
<tr>
<td>2006</td>
<td>terror, saved, intruder, energy, user</td>
</tr>
<tr>
<td>2007</td>
<td>badger, since, high, turk, turkey</td>
</tr>
<tr>
<td>2008</td>
<td>crew, speech, warning, saved, team</td>
</tr>
<tr>
<td>2009</td>
<td>knowledge, acker, team, album, country</td>
</tr>
<tr>
<td>2010</td>
<td>posted, member, protocol, kernel, security</td>
</tr>
<tr>
<td>2011</td>
<td>contact, security, village, holding, highlander</td>
</tr>
<tr>
<td>2012</td>
<td>saved, contact, team, underground</td>
</tr>
<tr>
<td>2013</td>
<td>team, forgive, security</td>
</tr>
<tr>
<td>2014</td>
<td>eagle, crew, electronic</td>
</tr>
<tr>
<td>2015</td>
<td>clash, king, terrorism, visit, alligator</td>
</tr>
<tr>
<td>2016</td>
<td>marocain, turk, steel, anonymous, team</td>
</tr>
</tbody>
</table>
Adoption of Malicious Content

The graph shows the adoption of malicious content from 1998 to 2016. The red bars represent the percentage of malicious content, while the red line indicates the trend. The adoption rate has significantly increased over the years, reaching a peak in 2014 and showing a substantial trend towards 2016.

Key points:
- The percentage of malicious content started at a very low level in 1998.
- There was a steady increase in adoption until around 2010.
- A significant spike occurred in 2014, indicating a rapid rise in malicious content.
- The trend line suggests a continued increase in adoption, although the rate of increase has slowed compared to previous years.

This data highlights the growing threat of malicious content in digital environments.
Adoption of email & Twitter handlers
General Trends

Key Observations
Teams vs. Individuals

#OpFrance

Owned!

U Remember Us ? We Are Back Again!

I Challenge Your Government To Stop Us!

We are: # Dr.Raw # Anonymoose # Maxtremen'C: | stars # LogChat Me # /SINICTI # CyberHacker Al

#OpFrance

France Coupes Website Hacked

We Are Not Terrorists, And Stop Calling Terrorist
Respect Existence, Or Expect Resistance
We Love Peace

I Challenge Your Government To Stop Us!

We are: # Dr.Raw # LogChat Me # Svennad Attackers #4 prudence/islam
Templates & Campaigns

Template

Hacked by Anonymous R4BIA Team
https://www.facebook.com/R4BIAHACKERSTEAM

please admin or visitor I believe your whole life will change after seeing these videos
if you are! the fourth video :) you must see it

Translation in the top of the page at your left hand

what will happen to you after your death!
why you are exist? What is your destiny?

Customization

Hacked by Freedom

please admin or visitor I believe your whole life will change after seeing these videos

Are you exited?
Ask yourself
what will happen to me after my death!
Who am I?
why I exist?
What is my destiny?
These questions and made the whole world.


copy & paste to start your own campaign.
Key Observations

1. Actors **cooperate** in teams
   Especially if driven by **similar, strong ideologies**

2. Defacements are organized around **campaigns**

3. Teams re-use a common **template** that each member can customize
Key Observations

Next Generation*
Defacement Explorer

(DefPloreX-NG)

(*) 1st generation presented at BH Arsenal
Deface Page Analysis

Defacement pages

De-duplication (hash based)

Metadata (e.g., timestamp, UF)
Campaign Detection

- Defacement pages
  - De-duplication (hash based)
- Deface Page Analysis
  - Dynamic analysis
  - Static analysis
  - Extracted content
- Metadata (e.g., timestamp, URL, reporting source)
- Campaign Detection
  - Normalization
  - Feature Extraction
  - Elastic
### Pre-Filtering Search

- **Deepwater Horizon**

### Campaigns Search

- **params.cluster_size > 1**

#### Choose clustering features

<table>
<thead>
<tr>
<th>Size</th>
<th>Key</th>
<th>Start</th>
<th>End</th>
<th>N. Attackers</th>
<th>N. Text Hashes</th>
<th>N. TLDs</th>
<th>N. Domains</th>
<th>Timeline</th>
</tr>
</thead>
<tbody>
<tr>
<td>72</td>
<td>🇪🇺</td>
<td>2003-04-01</td>
<td>2011-08-01</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>56</td>
<td>🇪🇺</td>
<td>2001-05-01</td>
<td>2016-09-01</td>
<td>42</td>
<td>4</td>
<td>13</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>🇪🇺</td>
<td>2011-05-01</td>
<td>2011-05-01</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>🇪🇺</td>
<td>2016-08-01</td>
<td>2016-08-01</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>🇪🇺</td>
<td>2011-05-01</td>
<td>2011-05-01</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>28</td>
<td></td>
</tr>
</tbody>
</table>
Implementation Details
Feature Engineering

Format of Title

Visual Features (colors)

Social Features

Visual Features (images)

Structural Features
Feature Engineering

Email Addresses

Media resources and URLs
# Summary of Features

<table>
<thead>
<tr>
<th>Group</th>
<th>Feature Name</th>
<th>Type and range</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual</td>
<td>No. of Images</td>
<td>integer ([0, \infty)]</td>
<td>Number of <code>&lt;img&gt;</code> tags</td>
</tr>
<tr>
<td></td>
<td>Perceptual Hash</td>
<td>binary (64 bits)</td>
<td>Calculated on the north-centered 1600x900 screenshot crop</td>
</tr>
<tr>
<td></td>
<td>Average Color</td>
<td>3 floats (RGB)</td>
<td>Average of the 5 most common colors in the screenshot</td>
</tr>
<tr>
<td></td>
<td>No. of Sound URLs</td>
<td>integer ([0, \infty)]</td>
<td>Number of URLs pointing to sound-hosting services or files</td>
</tr>
<tr>
<td></td>
<td>Type of first Sound URL</td>
<td>categorical</td>
<td>File and service type of the first (usually the only) sound URL</td>
</tr>
<tr>
<td>Structural</td>
<td>No. of each <code>&lt;tag&gt;</code></td>
<td>7 integers ([0, \infty)]</td>
<td>Number of style, embed, script, meta, object, iframe, and <code>a</code> tags</td>
</tr>
<tr>
<td>Geographical</td>
<td>Encoding</td>
<td>categorical</td>
<td>Detected text encoding</td>
</tr>
<tr>
<td></td>
<td>Language</td>
<td>categorical</td>
<td>Detected language (for labeling only)</td>
</tr>
<tr>
<td>Domains</td>
<td>External domains</td>
<td>real ([0, \infty)]</td>
<td>Ratio of links pointing to cross-origin domains</td>
</tr>
<tr>
<td></td>
<td>Letters in external domains</td>
<td>real ([0, \infty)]</td>
<td>Avg. ASCII letters in the external domains string</td>
</tr>
<tr>
<td>Social</td>
<td>No. of online handlers</td>
<td>int ([0, \infty)]</td>
<td>Twitter @handlers, hashtags, e-mail addresses</td>
</tr>
<tr>
<td>Title</td>
<td>Letters, digits, punctuation,</td>
<td>4 real ([0, \infty)]</td>
<td>Ratio of the listed character classes in the page title</td>
</tr>
<tr>
<td></td>
<td>white-spaces in title</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Campaign Detection: Clustering

- **BIRCH**
  - Balanced Iterative Reducing and Clustering Hierarchies

- **Do not** materialize the entire distance matrix
  - Statistical values are efficient to compute
  - Quickly find the closest cluster for each new data point
Scaling Clustering

- Scalability of BIRCH vs. DBSCAN (10 runs)
Clusters → Campaigns: Labeling

- Each cluster represented in a **succinct report**
  - **Time** span
  - Screenshot **thumbnails** (by perceptual hash)
  - Name of **actors** and **teams**
  - **Keywords** used in campaigns (e.g. #opfrance)
  - **Category** of targets (e.g, news, governmental sites)
Cluster 4

- portfolio (0.524)
- nicholas (0.355)
- steel (0.355)
- close (0.254)
- mohan (0.186)

- your (0.531)
- cyber71 (0.214)
- lucif42 (0.214)
- 5pid32 (0.214)
- that (0.212)

- kurdish (0.236)
- saved (0.236)
- kährd (0.236)
- tährk (0.236)
- turkish (0.236)
Findings
43% of actors join one team, at least

Half of joint campaigns are larger than 3

70% of campaigns are collaborative
Geopolitical Real-World Events

Successfully detected

- Sep 2001: 34 clusters, 621 pages
- Feb 2002: Gujarat Riots, 7 clusters, 7 pages
- Dec 2006: 14 clusters, 214 pages
- Jun 2009: Death of M. Jackson, 4 clusters, 79 pages
- Apr 2010: Deepwater Horizon Oil Spill, 5 clusters, 22 pages
- May 2011: 204 clusters, 471 pages
- 2012–2016: Battle of Aleppo, 259 clusters, 668 pages
- Nov 2014: Junker President of EU, 1 cluster, 1 page
- Jan 2015: Charlie Hebdo Shooting, 204 clusters, 664 pages
- Nov 2016: D. Trump President, 8 clusters, 22 pages
The “Charlie Hebdo” case

Hello.. Admin
I'm not Charlie
We are not Charlie
This message is clear
We are here to protect our prophet Mohamed
#OpFrance
#OpCharlie
Beware the hybrids, Doctor Ym
The “Charlie Hebdo” case

- Campaigns
- Teams
- Actors
The “Charlie Hebdo” case
Joint campaigns conducted by *anonymous*-affiliated groups against governmental sites.
Long-Term vs. Aggressive Campaigns
Long-Term vs. Aggressive Campaigns

Spot, aggressive campaign in reaction to war events
Example of a **long-running campaign** named *h4ck3rsbr*. The horizontal bars represent the different sub-campaigns (60) with their most targeted TLDs and teams (next slide). This is a very generic campaign, with different affiliates.
Example of a long-running campaign named h4ck3rsbr. The horizontal bars represent the different sub-campaigns (60) with their most targeted TLDs and teams (next slide). This is a very generic campaign, with different affiliates.
Conclusions

- Dark propaganda
- Prevailing phenomenon
- Driven by geopolitical motivations
- Key targets, influencing sites
References

- DefPloreX-NG
- GitHub: https://github.com/trendmicro/defplorex

(*) Joint work with Marco Balduzzi, Ryan Flores, Lion Gu and Vincenzo Ciancaglini
Using Machine-Learning to Investigate Web Campaigns at Large

Dr. Federico Maggi, @phretor

Joint work with Marco Balduzzi, Ryan Flores, Lion Gu and Vincenzo Ciancaglini