
URL Block List 2.0 - Expanding Data Sets for Precision Reporting and Deeper Data Mining

Patrick Cain

pcain@apwg.org

pcain@coopercain.com



Unifying the
Global Response
to Cybercrime

Overview

- The APWG has been accepting phishing URLs for a long time
 - Expanded to accept 'domains' a few years ago
- Metrics and reports are generated from this data
- But 'URL' doesn't give us:
 - DNS Data
 - IP Address
 - Actual lure or message
- Can we expand the UBL to include more data?
 - Email addresses, fone #s, ?, ?



Unifying the
Global Response
to Cybercrime

The Processing challenge

- We accept data via email, or email, or email.
 - This may not work with an expanded data set
- Divided submitters into low-volume (LV) and high-volume (HV)
 - Devise a web page for the LV
 - Develop scripts for the HV
- You get the data via a 'pull' from a repository
- Hard to search – manually or scriptily



Unifying the
Global Response
to Cybercrime

Preparing for the Future

- As phishing changes you've asked for more data types
 - Phishkits
 - Infected systems
 - Malware distributors
 - Attacking systems (?)
- Can we use the same processing for these new things?



Unifying the
Global Response
to Cybercrime

The webpage (draft)

APWG Committed to Wiping Out Internet Scams and Fraud

Current Display Language: en-US
version: 1.2

Submit Events to Our Data Clearinghouse (PILOT ACTIVITY)

Start Here:

- About This Effort
- Support
- Testing?
- Report a Bug

Welcome to the APG eCrime Reporting Pilot.

Please use this form to report eCrime activities to the APWG, which will add malicious URLs and IPs to its blocking lists and aggregate and forward relevant information to others.

We ask for significantly detailed information in this form. Although you could just provide the required data -- as that would reduce the number of eCrime victims -- the additional data submitted allows researchers and investigators to correlate activities to identify and actually catch the 'bad guy'.

Please select the language to use to display text fields:

Submitting data is performed in three steps: 1) Enter your contact data; 2) Enter the source and destination information for the suspicious activity; and 3) tell us the sensitivity and redistribution restrictions on your submission.

(Step 1 of 3): Your Contact Information >>

About the APWG Privacy Policy Contact Us Copyright ©2012 Pat Cain., All rights reserved.

Taskbar icons: Windows, Internet Explorer, File Explorer, Notepad, Computer, Firefox, two instances of a blocked icon (X), and a yellow icon.

eCrime '12
October 22-25
Puerto Rico
APWG



Unifying the
Global Response
to Cybercrime

Data selector



Committed to Wiping Out
Internet Scams and Fraud

Current Display Language: en-US
version: 1.2

Submit Events to Our Data Clearinghouse (PILOT ACTIVITY)

Start Here:

About This Effort

Support

Testing?

Report a Bug

(Step 2 of 3): Event Details

What Type of Incident was this?

Please select a crime type

Please select a crime type

Phishing for Credentials

Mule Recruitment

Mule Activity

Bot or Infected System Activity

Unknown

[About the APWG](#)

[Privacy Policy](#)

[Contact Us](#)

Copyright ©2012 Pat Cain., All rights reserved.



Unifying the
Global Response
to Cybercrime

Phish page

Submit Events to Our Data Clearinghouse (PILOT ACTIVITY)

(Step 2 of 3): Event Details

What Type of Incident was this?

Phishing for Credentials

Report a Phishing Event

(version: 1.2)

What type of fraud did the lure/message attempt?

credential phishing - The lure asked for your cre *

The Name of the Organization Phished:

How did you receive this phishing lure?

email message

What is your confidence that this is a real phishing site?

7% - This is a test submission. Do not include irj *

The number of these lures received:

0

Please paste the *FULL* contents -- including headers -- of the received lure below.
We will try and decode relevant data to save you typing.



APWG

Unifying the
Global Response
to Cybercrime

Bot Page

Submit Events to Our Data Clearinghouse (PILOT ACTIVITY)

(Step 2 of 3): Event Details

What Type of Incident was this?

Report Bot Infection or Activity

(version 1.1)

What Date and Time was the Infection or Activity Seen? GMT

What is the Infected System's IP Address? * v4

What is the Infected System's DNS Name?

How confident are you of this infection? *

The system is infected with what Bot or Infection (e.g., Zeus, Cutwail)?

This System was a member of which botnet?

The system was performing this activity (select or enter text)

How was the infected system detected?
 Traffic Analysis
 Check-into C & C
 Contacted a Sinkhole

What was the Command and Control Systems's IP Address? v4

What is the Command and Control's DNS Name?

Status

- The web pages are at repoman.apwg.org/forms
- The scripts are on our ecrisp-x page
 - Sourceforge.net/projects/e-crispx
- There is a trac instance on repoman to report (the one last) bugs
- We currently accept phish and botz
 - Phish → UBL
 - Botz →
 - A search and retrieve interface is coming



Unifying the
Global Response
to Cybercrime

Related work: How to convey policy/sharing info

- Restriction markings
 - How to mark: Share with LEO? Friends? Public?
 - How to show: Know but no Touch!
- Can this data be shared with law enforcement?
 - There may be legal or other reasons involved
- Has the data aged out of sensitivity?
 - An IP with a bot in 1994 is not secret in 2012!
- We are piloting splitting the 'marking':
 - What is the sensitivity of the data
 - How we can share it



Unifying the
Global Response
to Cybercrime

silly and real examples

- Sharing with the 'Public':
 - 0 - Do not share
 - 1 - Summary data may be shared
 - 2 - Details may be shared
 - 3 - Too late. (already shared)
- Sensitivity marking:
 - Historical - Data is probably already known or is relatively old
 - Live-overt - Data was collected in a known fashion or collector
 - Live-covert - Data was collected via a private collection facility
 - Active - Data is currently part of an active LEO investigation
 - Active-no touch - Do not disturb or take actions on this dataset
 - Unspecified or Unknown



Unifying the
Global Response
to Cybercrime

An example...

- How can this data be shared within the APWG/xxx?
 - 0 - For recipient use only. Not to be shared at all.
 - 1 - Recipient(s) should NOT share details of this data outside of group members
 - 11 - Recipient(s) may share summary data with their internal groups
 - 13 - Recipient(s) may share details with their internal groups
 - 21 - Summary data may be shared with other trusted security types
 - 23 - Details may be shared with other trusted security types
 - 31 - Summary data can be shared freely with the public
 - 33 - Details can be shared freely with the public
 - 25 - Shared with black hats only!



Unifying the
Global Response
to Cybercrime

Restrictions Pilot

mitted to Wiping Out
let Scams and Fraud

Current Display Language: en-US
version: 1.2

Submit Events to Our Data Clearinghouse (PILOT ACTIVITY)

(Step 3 of 3): Enter the last few details

A fantastic use of collected event data is sharing it within a group, law enforcement, or other security researchers. In this section you tell us where we can share this data.

What is the default sensitivity of this report?

Unspecified or Unknown

How can this data be shared? [Selecting a value implies lesser values.]

23 - Details may be shared with other trusted se

Error Messages
error-type
resource-uri
response-status-code
response-headers
response-reason-phrase
response-body
stuff:

- 0 - For recipient use only. Not to be shared at all.
- 1 - Recipient(s) should NOT share details of this data outside of members
- 11 - Recipient(s) may share summary data with their internal groups
- 13 - Recipient(s) may share details with their internal groups
- 21 Summary data may be shared with other trusted security types
- 23 - Details may be shared with other trusted security types
- 31 - Summary data can be shared freely with the public
- 33 - Details can be shared freely with the public
- 25 - Shared with black hats only!

SUBMIT >>

Open excitement

- Forms Integration into ECX
- How to integrate ECX data
 - Data from ECX entries into distributions
 - Script data into ECX
- Retrieval



Unifying the
Global Response
to Cybercrime

Your part...

- Feel free to submit something
 - There is a 'confidence' for "test" = not distributed
 - There is a password on the web page
 - user=dataTester pw=apwg2012 *
- The interface from the pages to the UBL is a little slow right now & the search is disabled.
- If you have ideas on how to give you the data, please advise.

- The repository maillist may become active again



Unifying the
Global Response
to Cybercrime



Thank You

pcain@apwg.org

[Send us data in IODEF.]

[Wait. Send us BOT data in IODEF.]



Unifying the
Global Response
to Cybercrime