

Discovering Phishing Dropboxes Using Email Metadata

Tyler Moore¹ and Richard Clayton²

Computer Science & Engineering Department,
Southern Methodist University, Dallas, TX¹
Computer Laboratory, University of Cambridge, Cambridge, UK²

APWG eCrime Researchers Summit
October 23, 2012



Outline

- 1 Introduction and research approach
 - Phishing kits
 - Incoming email metadata
- 2 Techniques for identifying dropboxes
 - Direct identification of dropboxes
 - Indirect identification of dropboxes
 - Identifying the source of dropbox email



Phishing kits

The screenshot shows a web browser window displaying a ZDNet article. The browser's address bar shows the URL: www.zdnet.com/blog/security/diy-phishing-kits-introducing-new-features/1104. The article title is "DIY phishing kits introducing new features". The author is Dancho Danchev, dated May 15, 2008. The article text discusses the increase in phishing attacks and the introduction of new features in DIY phishing kits. A sidebar on the left contains social media sharing options (14 Comments, 0 Votes, Like, 0, Tweet, Share). A right sidebar features an IBM advertisement and a "Related Stories" section with links to articles about WebSense, BitDefender, Chinese hackers, and Kaspersky Labs. At the bottom right, there is a "ZDNet Newsletters" sign-up form.

DIY phishing kits introducing new features

Summary: What are some of the main factors for the increase of phishing attacks, and their maturity from passive emails to blended threats attempting to not just steal personal information, but also infect with malware by embedding client-side vulnerabilities at the pages? It's all a matter of perspective, which in this post will emphasize on the continuing efforts on behalf of phishers to innovate, and introduce new features within the most recently obtained do-it-yourself phishing page generators.

By Dancho Danchev for Zero Day | May 15, 2008 -- 08:02 GMT (01:02 PDT)
Follow @danchodanchev

Phishers:

- AIM
- Amazon
- ADL
- Balbo
- Chase Bank
- Citi Bank
- Click And Buy
- Ebay
- Facebook
- File Front
- Friendster
- Game Battles
- Gmail
- Hotmail
- ICQ
- iTunes
- Money Bookers
- Myspace
- Nexon
- Paypal
- Photobucket
- Rapidshare
- Ripway
- Runescape
- Skype
- Xbox
- Yahoo Mail
- Youtube

FTP Phisher Directly to server

Host:

User:

Pass:

Path:

Related Stories

- Phishers are becoming smarter, more targeted: WebSense
- BitDefender releases tool for removing Gauss financial malware
- Chinese hackers linked to Canada's Telvent breach
- Help us crack Gauss' encryption: Kaspersky Labs

The best of ZDNet, delivered

ZDNet Newsletters

Get the best of ZDNet delivered straight to your inbox

ZDNet Must Read News Alerts - US: Major news is breaking. Are you ready? This newsletter has only the most important tech news nothing else.

Phishing kits: typical PHP code

```
<?php
$ip = getenv("REMOTE_ADDR");

$mess = "Email: " . $_POST['email'] . "\n";
$mess .= "PWord: " . $_POST['passwd'] . "\n";
$mess .= "IP: " . $ip . "\n";

$dest = "dropbox@example.com";
$subj = "PP ReZuLtZ";

if (mail($dest, $subj, $mess))
    { header("Location: /www.paypal.com/"); }
else
    { echo "ERROR! Please go back retry."; }
?>
```



The static nature of phishing kits

- PHP script invariably included in the ZIP archive
- Almost never edited on the server itself
- Thus it is inconvenient to change the subject line or dropbox email address as criminals move across servers
- This means we can link criminal behavior over time



Incoming email metadata

Timestamp

The time that the email is placed into a mailbox.

Source IP address

The machine that sent the email to the email provider.

SMTP “mail from”

The sender of the email, as declared in the SMTP conversation. This can be forged but usually provides some identification of true origin.

SMTP “mail to”

The destination(s) to which the email is being sent. In this context, this information is always valid.

From

‘From:’ email header field. It can be set by the phishing kit and is usually entirely bogus.

Subject

‘Subject:’ email header field (invariably set by phishing kit).

URLs

These are the URLs from the body of the email.



Why include URLs in email metadata?

- URLs are a very distinctive way to identify email spam
- Email addresses are treated as `mailto://` URLs
- If a phishing victim's account is an email address, then this will turn up in the metadata of a dropbox email



Finding dropboxes

How phishers use dropboxes:

victim1@example.com



Hacked Server

To: dropbox1@example.com
Subject: P1 ReZuLtUS
user: victim1@example.com
pass: hamster34

Dropbox

How we identify dropboxes:

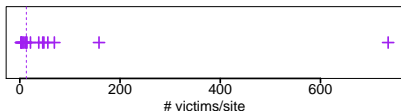
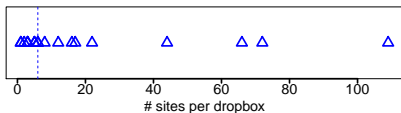
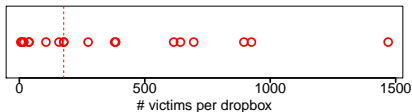


Direct dropbox discovery

- Our dataset
 - Phishing URL source: cleaned amalgamation of APWG, PhishTank, brand owner, and brand protection company feeds
 - On June 1 2012 we sent emails with spurious credentials to 170 different websites targeting PayPal reported in May 2012 and found to still be online
- Results
 - 28 / 170 emails found in email metadata logs (16.5%)
 - 17 distinct dropbox email addresses
 - Lots of distinctive Subject lines:
 - P1 ReZuLtUS
 - Paypal Spam Result
 - 10.0.0.1 | New PayPal Account
 - [EMAIL: jim@example.com | secret]



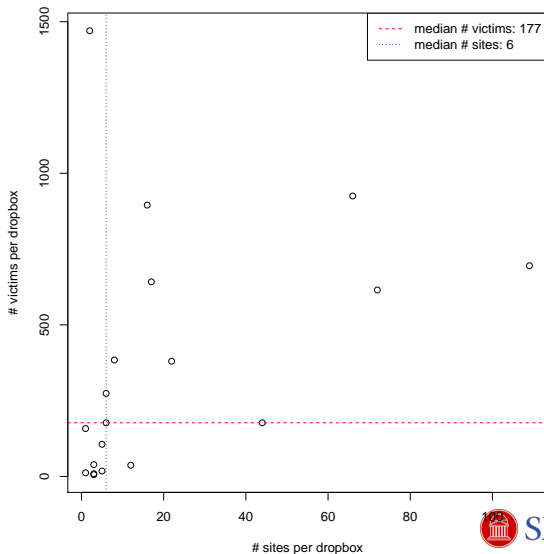
Observed dropbox statistics



	mean	median
victims (emails)	380	177
sites (senders)	22	6
<u>victims</u> <u>sites</u>	68.9	13



Limited correlation between # victims and # sites



Identifying more dropboxes by matching dropbox subjects

- Recall those distinctive subject lines?

P1 ReZuLtUS

Paypal Spam Result

10.0.0.1 | New PayPal Account

[EMAIL: jim@example.com | secret]

- We found 15 distinct patterns from the 28 subject lines
- Searching for those subject lines in all email yielded 81 new dropboxes (3 times as many as found through direct probing)



An upper bound on criminals targeting PayPal

- $\#$ dropboxes \sim $\#$ criminals
- Some criminals use multiple dropboxes, and criminals inevitably register new dropboxes as their old ones are shut
- But for a small snapshot in time, the $\#$ of dropboxes can serve as an upper bound for the number of criminals operating
- We found 29 dropboxes used to attack PayPal in July 2012 (17 throughout month and 12 for shorter periods)
- Thus we estimate that we found between 20-29 criminals, and our direct identification technique found dropboxes for 16.4% of PayPal phish
- So we estimate 122–164 criminals attacked PayPal in July 2012 (out of 26 900 distinct URLs on 13 018 domain names)



Intersection method to identify dropbox source URLs

1 Identify dropboxes from subject patterns

P1 ReZultUS	dropbox1@example.com
	dropbox2@example.com
Paypal Spam Result	dropbox3@example.com

2 Find victims from mailto: URLs in dropbox emails

dropbox1@example.com	time email received	victim email
V1	2012-06-08 01:28:10	mailto:victim1@example.com
V2	2012-06-08 21:00:01	mailto:victim2@example.com

3 Find phishing URLs by intersecting URLs in victim emails

V1 mailto URLs	V2 mailto URLs
amazon.com	nytimes.com
twitter.com	facebook.com
http://surses-paypal.com-confirm-cgi.bin. account-15f2vb1n.save-data-supportteam165 fgg478521fdds5ds1d6.dnstour.com/Uid=98635/	http://surses-paypal.com-confirm-cgi.bin. account-15f2vb1n.save-data-supportteam165 fgg478521fdds5ds1d6.dnstour.com/Uid=98635/



Interventions possible using the intersection method

- 1 Identify phishing victims at the time of credential disclosure
 - Regularly run searches for known dropbox subjects and identify victims from `mailto:` URLs
- 2 Identify (and block) phishing URLs faster
 - Once two victims have entered their details, identify the phishing URL and email provider can block its other customers from being phished
 - Can also pass along newly discovered URLs to blacklists



Intersection method proof-of-concept

- Inspected one week's worth of email for dropbox subjects (15-21 July)
- Found 934 victim credentials sent to dropbox from 114 IP addresses
- Of these, 159 victims had email address with metadata we could inspect coming from 47 IP addresses
- Of 47 IP addresses, 25 had one victim with metadata
- This leaves 22 potential phishing URLs to run intersection method



Intersection method in action

2012-07-19 15:16:22	phish arrived at V1
2012-07-19 15:20:02	phish arrived at V3
2012-07-19 15:21:32	V1 becomes a victim
2012-07-19 15:48:30	V6← http://77kids.com etc.
2012-07-19 16:12:56	V7← http://ui.constantcontact.com/...
2012-07-19 16:16:18	phish arrived at V5
2012-07-19 16:18:53	phish arrived at V4
2012-07-19 16:23:40	phish arrived at V2
2012-07-19 16:36:11	V2 becomes a victim
2012-07-19 16:37:25	V6← http://www.constantcontact.com
2012-07-19 16:39:16	V3 becomes a victim
2012-07-19 16:46:52	V4 becomes a victim
2012-07-19 17:13:02	phish arrived at V6
2012-07-19 17:32:48	V5 becomes a victim
2012-07-19 18:19:15	V6 becomes a victim



URL blacklists could benefit from intersection method

Phish	Time found by intersection	Time in blacklist	Lag
PHISH 3	2012-07-15 16:35:19	2012-07-02 21:27:12	–
PHISH 2	2012-07-16 23:01:02	2012-07-17 02:18:15	3.2 hrs
PHISH 4	2012-07-17 00:15:27	2012-07-21 11:13:06	4.5 days
PHISH 5	2012-07-17 01:13:40	2012-07-15 15:10:07	–
PHISH 6	2012-07-18 03:58:25	2012-07-18 06:21:28	2.5 hrs
PHISH 7	2012-07-18 18:54:24	2012-07-23 14:18:38	4.8 days
PHISH 8	2012-07-19 04:49:26	2012-05-16 18:37:49	–
PHISH 1	2012-07-19 16:36:11	never reported	∞
PHISH 9	2012-07-20 13:35:24	2012-07-17 20:11:35	–
PHISH10	2012-07-21 13:17:48	2012-07-18 00:05:03	–
PHISH11	2012-07-22 05:20:09	2012-07-20 14:28:44	–



Conclusion

- Dropbox email accounts are a critical but often overlooked component to most successful phishing attacks
- We have presented low-cost mechanisms to identify dropboxes by combining phishing URL lists with email metadata
- The techniques could be used to protect users and identify more phishing sites faster
- We estimate that 120–160 criminals targeted PayPal in July 2012 using 26 900 distinct URLs
- Increased attention to dropboxes could have a disruptive effect
- For more: <http://lyle.smu.edu/~tylerm/>

