SANS Institute
InfoSec Reading Room

Zombie profiling with SMTP greylisting

Email is consistently used to propagate malware, conduct phishing and deliver spam. A large proportion of this unwanted email is sent by compromised machines or computer zombies. This paper observes that computer zombies react differently to being greylisted, providing a method to profile computer zombies into various types. The GCIH course touches on the concept that this age is the <q>age of the botnets</q> and how malware is propagated with the help of email. This paper extends this topic by analysi...

Zombie profiling with SMTP greylisting

GCIH Gold Certification

Author: Jeremy Koster, jeremy.koster@gmail.com

Adviser: Joey Niem

Accepted:

Outline

1. Abstract

Email is consistently used to propagate malware, conduct phishing and deliver spam. A large proportion of this unwanted email is sent by compromised machines or computer zombies. This paper observes that computer zombies react differently to being greylisted, providing a method to profile computer zombies into various types. The GCIH course touches on the concept that this age is the “age of the botnets” and how malware is propagated with the help of email. This paper extends this topic by analysing greylisting activity for the purposes of identifying computer zombies and exploring methods to reduce the unwanted email received from botnets.

2. Introduction

In recent years, Internet users have been subjected to large quantities of unwanted emails such as spam, emails containing viruses and phishing emails. It is estimated that 85% of all email is unwanted (Messaging Anti-Abuse Working Group (MAAWG), 2008).

A computer zombie is a personal computer that is under the control of another individual without the knowledge of the computer's owner (Computer Zombie, 2008). A popular method for sending unwanted emails is by using thousands of computer zombies, known as a botnet, under the control of a botnet operator. According to Marshal TRACE

Report, “The top seven spamming botnets are responsible for 90% of spam” (Marshal 2008, Page 3). Botnet operators use their botnets to send spam, phishing emails and viruses. In doing so they can benefit financially or increase the number of zombies under their control. A prosperous business can be made by controlling or leasing the botnets to conduct click fraud, DDoS attacks, keylogging, distributing warez and sending unwanted email (Shadowserver 2007). Essentially, the larger the botnet and the more zombie computers under the control of the botnet operator, the greater opportunity the botnet operator has to gain financially. Without the capacity to send email successfully, the botnet owner can no longer grow their botnet and has less opportunity for financial gain by sending spam or phishing emails.

There are many methods employed to reduce the amount of unwanted emails received by the user. One of these methods designed to reduce unwanted emails from computer zombies is greylisting (Greylisting, 2008). Greylisting operates by issuing a temporary failure to email senders who have not previously sent email to the intended recipient. Unfamiliar senders are requested to re-send their email. A legitimate email server will typically recover well by re-sending the message quickly with no perceived impact to the end user. This is effective at reducing unwanted email because computer zombies have a limited capacity to correctly re-send failed messages. It has been observed that computer zombies continue to

incorrectly resend email, potentially giving away their intentions of sending unwanted email. This behaviour of incorrectly resending emails is the basis of research for this paper as a means of reliably identifying computer zombies and reducing their capacity to send unwanted email. It is not expected that a panacea for unwanted emails is discovered, but rather to identify techniques to raise the bar of email security and reduce the viability of email as a distribution method for botnet operators.

The SANS GIAC Certified Incident Handler (GCIH) course teaches the phases of handling an incident as Preparation, Identification, Containment, Eradication, Recovery and Lessons Learned. This paper focuses on the Preparation phase of incident handling as the methods researched may reduce the possibility of an incident occurring through the delivery of malicious emails. The GCIH course also discusses techniques to defend against worms and bots, both of which are commonly propagated by email. By researching methods of profiling computer zombies and ultimately blocking unwanted and malicious emails, it is hoped that techniques to defend against worms and bots will be expanded on.

3. Greylisting

Greylisting is a method that can be employed by a receiving email server to reduce the

amount of unwanted email. The basic concept is that on first receipt of an email from an unknown sender, the receiving server issues a temporary failure message to the sending party. The sending party, if a legitimate mail transfer agent (MTA), will attempt to send the same mail again within a few minutes (Greylisting 2008). Computer zombies will typically not attempt to send the same email again. A simplified example of an MTA responding to greylisting and the way a computer zombie responds to greylisting is demonstrated below.

Legitimate MTA response to greylisting

```
Received from [127.0.0.1] by from-domain.com; , 10 Mar 2008 11:43:10 +0100
From "user@from-domain.com" <user@from-domain.com>
To "user@to-domain.com" <user@to-domain.com>
451 4.7.1 Service unavailable - try again later

Received from [127.0.0.1] by from-domain.com; , 10 Mar 2008 11:44:12 +0100
From "user@from-domain.com" <user@from-domain.com>
To "user@to-domain.com" <user@to-domain.com>
250 2.0.0 Ok: queued as 2B8D71B3C2
```

Computer zombie response to greylisting

```
Received from [#.#.148.4] by mail.feha-spa####.de; , 9 Mar 2008 12:16:37 +0100
From "Francesca English" <humaner17@feha-spa####.de>
To "user@to-domain.com" <user@to-domain.com>
451 4.7.1 Service unavailable - try again later
```

```
Received from [#.#.148.4] by mailse####.gbc.net; , 9 Mar 2008 12:16:48 +0100
From "Joyce Talley" <auntss3@####herblocher.de>
To "user@to-domain.com" <user@to-domain.com>
451 4.7.1 Service unavailable - try again later

Received from [#.#.148.4] by mshgw-####.msh.de; , 9 Mar 2008 12:17:57 +0100
From "Harrison Gutierrez" <biggestz8@####imedia.zgs.de>
To "user@to-domain.com" <user@to-domain.com>
451 4.7.1 Service unavailable - try again later
```

From the examples above, the legitimate MTA uses the same sending IP address, FROM address and TO address when it retries sending the email after the initial temporary failure. The legitimate MTA waits one minute before sending the second attempt. The combination of IP address, FROM address and TO address has been seen before so the email is allowed to be delivered.

The computer zombie attempts to send three emails within a minute all with different FROM addresses. Each time the computer zombie retries to send an email the combination of IP address, FROM address and TO address has never been seen before so is rejected each time.

A possible reason for the different response is the difference in functionality and purpose between legitimate MTAs and computer zombies. Legitimate MTAs such as Postfix and Sendmail employ email queues which give the MTA the ability to retry failed sending

attempts quickly, reliably and predictably. The main purpose of a computer zombie is to deliver as much unwanted email as possible in the shortest amount of time possible. It may not have the functionality or resources to keep track of failed sending attempts to retry sending them in an expected time frame.

To facilitate greylisting, a table is maintained consisting of sender IP address, FROM address and TO address. When receiving an email the greylisting mechanism will check the table for a corresponding combination of IP address, FROM address and TO address. If none is found, a temporary failure will be issued. If a corresponding entry is found, the email will be accepted. Analysis of greylisting shows that it can be an effective measure against unwanted email, achieving an effectiveness of 97% (Harris 2003).

ID	Sender IP address	Recipient IP address	Sender email address
1	1.2.3.4	user@to-domain1.com	user@from-domain2.com
2	1.2.3.5	User2@to-domain3.com	user2@from-domain4.com

Table 1. Typical greylisting table

4. Research infrastructure

A greylisting module was written in PERL for the purposes of this paper. The PERL

script ran as a milter (Vierling 2004) using the PMilter library (Vierling 2007). Unlike a traditional greylisting mechanism the module was based on checking senders against a table of only two metrics, the sending IP address and sender's email address domain name (FROM domain). This method was chosen because it has less chance of delaying legitimate email and was expected to be as effective as the traditional method of greylisting.

ID	Sender IP Address	From domain
1	1.2.3.4	from-domain.com
2	1.2.3.5	From-domain2.com

Table 2. Research infrastructure greylisting table

The module was implemented on two email servers running the Postfix email server (Venema 2009). The greylisting module stored greylisting activity in a MySQL database. Mail exchanger (MX) records were created for a domain dedicated to capturing spam that pointed to the two mail servers. Email addresses in the domain were published on various forums that were publicly available. These email addresses quickly started to receive unwanted emails to an average of over 600 sending attempts per day. Sending attempts of unwanted email from

zombie computers were collected from the 1st of February to the 31st of August 2008. In total, information was captured from over 110,000 connection attempts that were made from over 19,000 different hosts.

Metrics collected for Zombie profiling

With every sending attempt, the greylisting module collected metrics to allow the profiling of computer zombies into zombie types. An example of a connection attempt made by a computer zombie is dislpayed below.

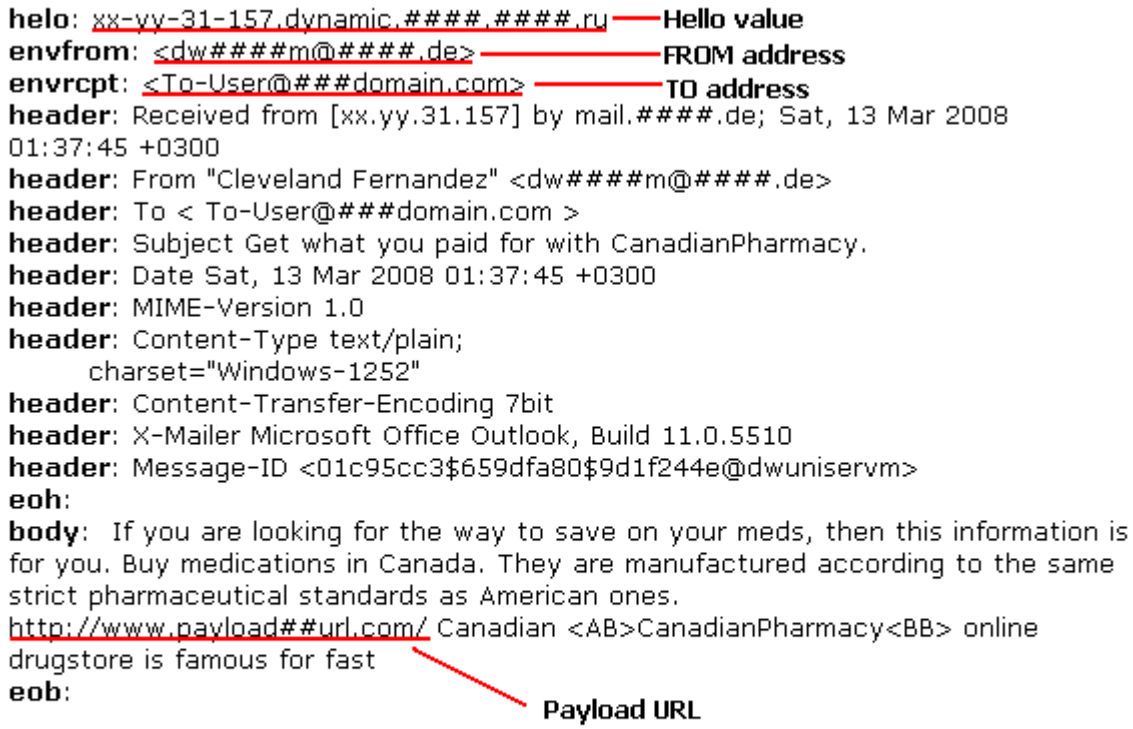


Figure 1. Connection attempt from computer zombie

IP address of sending host

This is the single IP address of each individual sending host. Each sending host was assigned a unique ID that is referenced in the following analysis. Every time a host with a new IP address made a connection, a new unique ID was allocated.

Hello value

As part of the SMTP protocol, an SMTP client must respond to the SMTP server's initial greeting with the Hello command of either ehlo or helo (Klensin 2001). The purpose of this is for the SMTP client to identify itself during client initiation. This Hello value can be seen as *xx-yy-31-157.dynamic.#####.#####.ru* in the example above.

Envelope "FROM" address (FROM)

The Envelope FROM address was collected during the first step of the SMTP mail transaction initiated with the "MAIL FROM:" SMTP command (Klensin 2001). This value consists of a user component and a domain component in the form of user@domain.com. The "FROM domain" component was used by the greylisting module to validate previous sending attempts. The FROM domain can be seen as *#####.de* in the example above.

Reverse DNS address of sending host

At every connection attempt the greylisting module will attempt to resolve the reverse DNS entry for the SMTP client by issuing a reverse DNS lookup (Reverse DNS lookup 2008). If a reverse DNS entry was successfully resolved this value was stored. If not, an "unknown"

entry was stored.

Timestamp of sending attempt (timing)

The date and time of each email sending attempt was logged at the time of initial connection.

Payload URL and Web server IP

Many unwanted emails will include a payload URL of a website to entice victims to buy a product, divulge personal information or infect their machine through downloading malware. The payload URL was collected as well as the IP address of the web server that the URL refers to. The payload URL is located in the body of the message and can be seen as *http://www.payload##url.com/* in the example above.

5. Zombie analysis

Using the metrics collected from the greylisting module the computer zombie behaviour was analysed. All connection attempts from 300 computer zombies included in the analysis, with computer zombies exhibiting similar behaviours grouped into computer zombie types.

95% of all sending hosts were able to be categorised. Of the 5% that were unable to be categorised, not enough information about the sending host was gathered by the greylisting module because only a single attempt at sending an email was made.

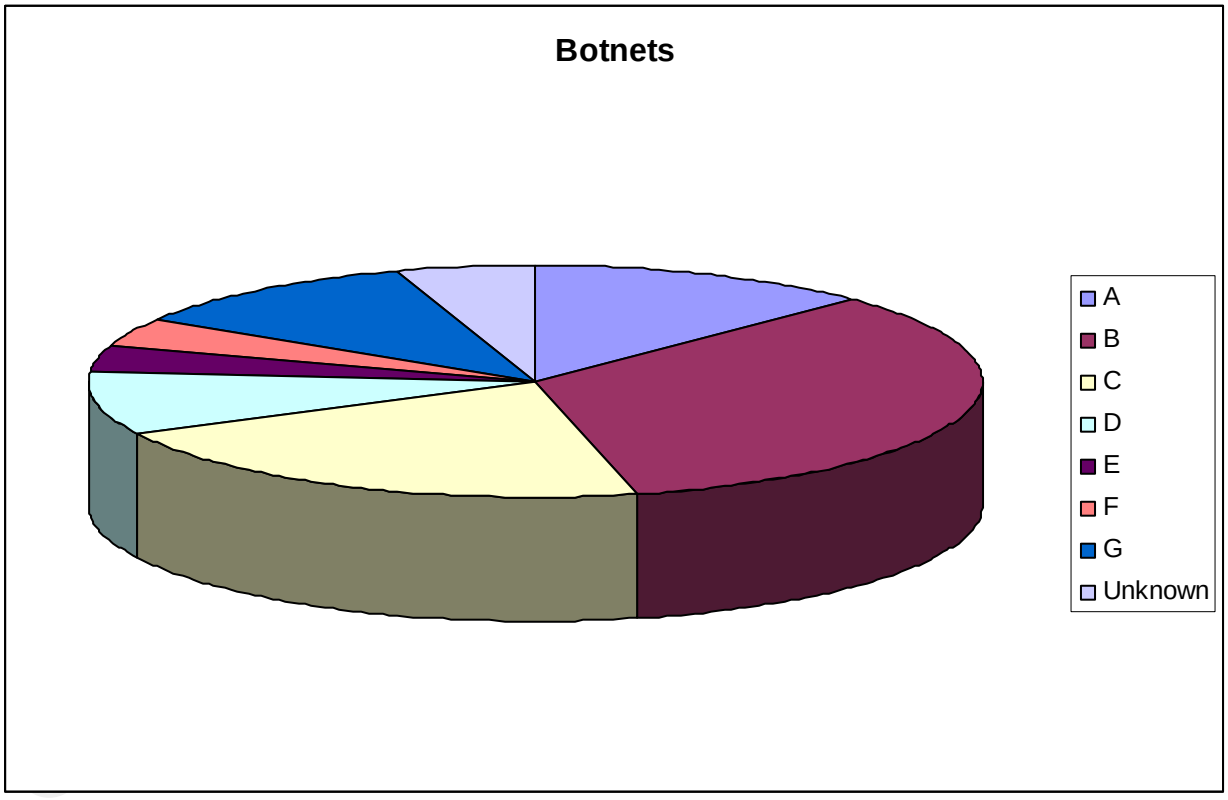
Type	Hello	FROM	Timing	Payload URL	Web server IP
A	no pattern.	FROM user is one word with 0-3 numbers at the end.	3 attempts in 1 minute.	Second level domain URLS are distinct always has www.*.com.	Same web server IP.
B	no pattern.	Always FROM domain is *.de May have lin*met in FROM user.	2-20 attempts in 1-17 minutes.	Uses the same URL for days. Sometimes includes random URL halfway through sending attempt.	Uses one URL to one web server IP.
C	no pattern.	FROM user is single word plus 0-3 random characters. FROM domain is commonly two words ".com".	3 attempts only.	Single host uses common URL. URL not common across multiple hosts.	Multiple URLS will use the same web server IP.
D	no pattern.	Uses same FROM user and FROM	3-4 attempts.	Commonly uses IP address as payload	Uses same web server IP for many days.

		domain during multiple attempts. Taken from sorted list FROM user begins with j[r/q].		Also includes legitimate URL in conjunction with payload URL e.g., bbc.co.uk.	
E	Payload URL used as hello value.	same FROM user and FROM domain used in each attempt. FROM domain is Payload URL.	1-2 attempts in 5 minutes.	Multiple URLs used.	Multiple URLS use the same web server IP addresses.
F	Uses reverse DNS if available otherwise IP address.	FROM user consist of "dws" %FROM_domain% "m". FROM domain starts with r,q,p or s	5 attempts in 5 minutes	Uses same URL for 3-5 days. Geocities pages used regularly.	Same web server IP address used.
G	Uses reverse DNS if available otherwise IP address.	Unusual characters in the FROM user value, numbers, underscores, hyphens and random capitalisation.	1 attempt only.	Uses same URL across zombie computers.	

Table 3. Computer zombie types and common behaviours.

Botnet Type	Count	Percent
A	37	12.3%
B	102	34%
C	63	21%
D	27	9%
E	10	3.3%
F	13	4.3%
G	33	11%
Unknown	15	5%
Total	**300**	**100%**

Table 4. Computer zombie distribution.



Graph 1. Computer zombie type distribution.

Behaviour 1. Resending emails as different sender

Botnet types A, B, C, E and F all resend emails using a different FROM user and FROM domain at each sending attempt. Botnet B is the most aggressive, at times trying up to 20 different FROM addresses in a few minutes.

Example 1.

Host id	Hello	FROM user	FROM domain	Timestamp	Payload	Web server IP
1034	221.%%.com.br	dodgingh	dimedis.de	18:16:10	2%%l.org	x.y.37.221
1034	221.%%.com.br	fittersmt8	work-os.de	18:16:27	2%%l.org	x.y.37.221
1034	221.%%.com.br	bossingmy	dropshop.de	18:16:46	2%%l.org	x.y.37.221
1034	221.%%.com.br	methodsq31	gemue.de	18:16:59	2%%l.org	x.y.37.221

In example 1, the sending host attempts to send email 5 times in under a minute all using different FROM domain. This computer zombie is an example of botnet type B.

Example 2.

Host id	Hello	FROM user	FROM domain	Timestamp	Payload	Web server IP
1035	feed%%g.com	Bidz.com	feed%%g.com	18:22:44	feed%%g.com	x.y.177.10
1035	feed%%g.com	iPhone	feed%%g.com	18:54:58	feed%%g.com	x.y.177.10

Example 2 shows where the same FROM domain is used in multiple sending attempts. It should also be noted that a different FROM user is used in each attempt. This computer zombie is an example of botnet type E.

Further analysis of the collected data reveals that out of 1,9013 hosts attempting to send unwanted email, 1,224 hosts used the same FROM domain when attempting to send multiple times. If this simple metric was used as a filter, 93.56% of sending hosts would not have been permitted to successfully deliver their email. It is expected that this behaviour would not normally be exhibited by a legitimate MTA and therefore may be useful in identifying zombie computers.

Behaviour 2. Common Payload URL and Web server

Botnets were observed to re-use a URL payload across multiple sending attempts. The computer zombies use the same payload URLs or web servers over and over again in an apparent attempt to brute force the email with payload into the recipients mailbox.

Example 3

Host id	FROM user	FROM domain	Date	Time	Payload
902	receding	vestolit.de	1/02/2008	12:43:03	home.gr%%iti.net

902	contoured86	maba.de	1/02/2008	12:43:27	home.gr%%iti.net
902	commonerslp378	sennebogen.de	1/02/2008	12:43:53	home.gr%%iti.net
902	pantheistsrur68	kronen-gmbh.de	1/02/2008	12:44:18	home.gr%%iti.net
902	monolithssa	loonataraxis.de	1/02/2008	12:44:44	home.gr%%iti.net
902	moundinglay9	pavillon-hannover.de	1/02/2008	12:45:09	home.gr%%iti.net
902	dibblessx	oberberg-online.de	1/02/2008	12:45:35	home.gr%%iti.net
902	dazzlesua	wcv-mail.de	1/02/2008	12:46:00	home.gr%%iti.net
902	distractionsi	drk.de	1/02/2008	12:46:26	home.gr%%iti.net
977	filets	nurnber-land.de	2/02/2008	4:10:31	2%%l.org
977	catatonicsl22	buecherungel.de	2/02/2008	4:10:41	2%%l.org
977	rilesye	hoeren.de	2/02/2008	4:10:53	2%%l.org
977	extremercibk	laqueur.de	2/02/2008	4:11:01	2%%l.org
977	ramrodwu	polymergmbh.de	2/02/2008	4:11:11	2%%l.org
1200	postcode25	lfl.bwl.de	3/02/2008	19:56:29	2%%l.org
1200	hitchcock	stadt-laage.de	3/02/2008	19:56:43	2%%l.org
1200	disenchanting	awatrade.de	3/02/2008	19:56:55	2%%l.org
1200	bumpkiny8	werbug.nvag.de	3/02/2008	19:57:16	2%%l.org
1200	syndionymx1	hymer-auer.de	3/02/2008	19:57:28	2%%l.org

In example 3, Botnet B used a single URL for all sending attempts for one day and then swapped to different URL for three days. It should be noted that this URL was not seen to be used by any other botnet.

Example 4

Host id	FROM user	FROM domain	Date	Time	Payload	Web server IP
1084	clunkier234	qtautonews.com	3/02/2008	1:53:12	www.pok%%fast.com	x.y.192.91
1084	whitsundayki9	nefcestest.com	3/02/2008	1:54:41	www.cb%%bxv.com	x.y.192.91
1084	shoddily4	iksny.com	3/02/2008	1:55:58	www.tif%%est.com	x.y.192.91
1092	contuousso866	playdium.com	3/02/2008	2:30:05	www.car%%ius.com	x.y.192.91
1092	worthier	gillbee.com	3/02/2008	2:31:15	www.pok%%fast.com	x.y.192.91
1092	shuttinge3	boat2view.com	3/02/2008	2:32:15	www.tob%%tobe.com	x.y.192.91

In example 4, Botnet A used 3 different URLs in a single round of sending attempts, but all 3 refer back to the same web server. It also appears the Botnets do not share payload URLs or web servers between botnet types.

Further analysis of the collected data reveals that out of 19,013 hosts attempting to send unwanted email, 17,264 included a URL payload in their email. 90% of all hosts use URL payloads. 4,304 unique URLs were used (second level domains) and 2,703 unique web servers were referenced. On average 7 hosts would reference a single web server.

The behaviour of re-using URL payloads and web servers may allow for the filtering of emails based on these metrics. If a botnet has a limited number of URLs that it uses as a payload then once the URL has been identified as being used by a botnet, future hosts using the URL can be identified as computer zombies. In short, a computer zombie can be reliably identified

by being associated with a payload. This can help in the distinguishing zombie computers from legitimate MTAs.

Behaviour 3. Single try

Botnet G attempts to send a single email, but never returns to try again. The simple act of issuing a temporary failure seems to confound the computer zombie causing is to never return.

Example 5

Host id	Hello	FROM domain	Date	Date	Payload	Web server IP
1093	y.x.dyn.%%.%%.com	richelle-shobana	eadnet.it	3/02/2008	2:53:03	sheet%%ade.com

Further analysis of the collected data reveals that out of 19,013 hosts attempting to send unwanted email, 3,272 hosts never return after the initial temporary failure. If this behaviour was used as an email filter 17% of hosts would have been unable to deliver their intended emails. Legitimate MTAs are expected to retry reliably after a temporary failure. Hosts that do not retry can be reliably identified as computers zombies.

Behaviour 4. Common entities in FROM addresses

Botnet types B, D and F all place common characters or patterns in the FROM address.

Example 6

Host id	FROM user	FROM domain	Timestamp	Payload	Web server IP
935	gridpj4	buw.de	19:46:41	home.graf%%.net	x.y.123.99
935	behalfiz1	vanillae.de	19:47:13	home.graf%%.net	x.y.123.99
935	extendingaza4	kumavision.de	19:47:35	home.graf%%.net	x.y.123.99
935	markupsbi9	huethig.de	19:47:52	home.graf%%.net	x.y.123.99
937	tennesseetmu5	ba-karlshe.de	19:55:47	home.graf%%.net	x.y.123.99
937	ifymjb	tiptop.de	19:57:18	home.graf%%.net	x.y.123.99
937	blacktoppedk73	oel-maier.de	19:57:56	home.graf%%.net	x.y.123.99
937	touslee4	audi-club.de	19:58:59	home.graf%%.net	x.y.123.99

In example 6, the FROM domain used by botnet B ended in "de". This was sustained over the 6 month collection period.

Example 7

Host id	FROM user	FROM domain	Date	Timestamp
1166	jradcliff	certifiedreports.com	3/02/2008	12:45:33
1166	jradcliff	certifiedreports.com	3/02/2008	12:45:51

Host id	FROM user	FROM domain	Date	Timestamp
1170	jquagliano	quagseeg.com	3/02/2008	13:36:01
1170	jquagliano	quagseeg.com	3/02/2008	13:36:03
1170	jquagliano	quagseeg.com	3/02/2008	13:36:05

From Example 7, sending attempts from botnet D have a FROM user that begins with “jq” or “jr”. This pattern is sustained over the 6 month collection period.

Exmaple 8

Host id	FROM user	FROM domain	Date	Time
1028	dwsigabm	sigab.ch	2/02/2008	17:45:23
1028	dwsjavarkjallarinnm	sjavarkjallarinn.is	2/02/2008	17:45:37
1028	dwsenefrom	senefro.org	2/02/2008	17:46:04
1028	dwsequislifem	sequislife.com	2/02/2008	17:46:17

From Example 8, sending attempts from botnet F have a FROM user pattern of “dws” FROM domain “m”. This pattern is sustained over the 6 month collection period.

Through the analysis of sending attempts it is clear that botnets can be categorised into types by the patterns and common characters in their attempts of sending unwanted emails. Although this is useful for the identifying types of botnets, it may not be a reliable method of

differentiating zombie computers from legitimate MTAs as legitimate MTAs are not required or expected to omit certain characters from FROM addresses. A combination of character patterns and failed attempts to resend email might be possible and may be worth further investigation beyond this paper.

Behaviour 5. Payload URL used in Hello and FROM address

Botnet type E uses a payload URL as the Hello command and the FROM Domain as seen in example 9 below.

Example 9

Host id	Hello	FROM user	FROM domain	Payload
1069	sl%%basketb%%l.com	lovematch	sl%%basketb%%l.com	sl%%basketb%%l.com

Even though this may be unusual behaviour for an MTA to exhibit, it is not entirely unexpected. As a method for identifying zombie computers this may not have a high reliability but might be useful in combination with other behaviours.

Changing botnets

It was observed that a few zombie computers changed behaviour during the collection period

which gave the impression that the zombie computer had moved botnets.

Example 10

Host id	Hello	From suer	FROM domain	Date	Time	Payload
1079	l%%s.com	jramirez	bppr.com	3/02/2008	1:23:24	www.cos%%.com
1079	l%%s.com	jramirez	bppr.com	3/02/2008	1:23:43	j%%ut.com
1079	l%%s.com	jramirez	bppr.com	3/02/2008	1:23:43	www.cos%%.com
1079	l%%s.com	jramirez	bppr.com	3/02/2008	1:23:52	j%%ut.com
1079	l%%s.com	jramirez	bppr.com	3/02/2008	1:23:52	www.cos%%.com
1079	x.y.25.61	dwsouthsescrowm	southsescrow.com	19/03/2008	20:18:04	www.se%%iv.com
1079	x.y.25.61	dwstalterim	stalteri.com	19/03/2008	20:18:36	www.se%%iv.com
1079	x.y.25.61	dwspecialtym	specialty.com	19/03/2008	20:18:46	www.se%%iv.com
1079	x.y.25.61	dwsmithscottm	smithscott.com	19/03/2008	20:18:56	www.se%%iv.com
1079	x.y.25.61	dwstagingm	staging.com	19/03/2008	20:19:05	www.se%%iv.com

In example 10, host 1079 displays behaviour in February 08 of belonging to botnet type D (FROM user begins with “jr” and legitimate URLs in payload). In March the host displays behaviour of belonging to botnet type E (FROM user comprises of “dws” FROM domain “m”). This swapping of behaviour could be for a number of reasons:

- The zombie software may be upgraded by the botnet operator.

- The vulnerable machine may be taken over by another botnet with new zombie software loaded.

- The IP address that the host is sending from may have multiple computers behind it in a network address translation (NAT) configuration.

This behaviour may be useful to link certain botnets together if an IP address is already seen to be sending unwanted emails from a zombie computer. Further information from this IP address can help to identify other zombie computers. This is again identifying computer zombies by association to another host's behaviour.

6. Possible applications and further research

Greylisting would normally be applied to the external MTA of an organisation to reduce the amount of inbound unwanted mail. A number of possible applications of the information gathered from greylisting and areas for further research are suggested below.

Blacklist targeted attacks

Organisations around the world are constantly bombarded with unwanted emails from computer zombies from all over the Internet. A particular protection mechanism is to employ an IP blacklist provided by a commercial organisation. Such a blacklist is provided at a fee to

allow a customer to block unwanted emails from hosts that are known to send unwanted email. While effective, this technique will only block the zombie computers that the commercial organisation is aware of. Computer zombies that are specifically targeting a particular organisation may not be identified and included in the IP blacklist.

By including data collected from greylisting on their own email servers, an organisation may build a blacklist of computer zombies that are specifically targeting only their own infrastructure. In conjunction with a commercial IP blacklist, this may offer a more complete protection against unwanted email.

Web filtering

During the collection of the research data a list of payload URLs located in unwanted emails was also gathered. As this list was comprised mostly of spam and phishing sites it would be useful to apply this list to an organisation's web proxy to assist in filtering out dangerous URLs. This could reduce the number of computers infected by malware caused by users visiting malicious websites.

IDS and event correlation

The data collected and the behaviours identified by zombie computers, could be included in

an IDS signature or event correlation software. The behaviours such as sending repeated emails with different FROM addresses or consistent URLs could be detected by an IDS and alert system administrators of botnet activity on the internal network. Additionally, a list of computer zombie IP addresses or DNS names in use by botnet operators could be used by event correlation software to alert staff when attempts to contact IP addresses and lookup DNS names are made. This could give early notification of botnet behaviour on the internal network.

Outbound email filtering

Phishing emails often contain an email address as a means for a victim to divulge personal information. Email addresses known to be included in the body of emails from zombie computers could be compiled into a list and used to block email being sent by users. This could protect an organisation's staff from sending personal information to malicious parties. This could also alert the Information Security team of staff who should be included in Information Security awareness programs.

Additional payloads and metrics

The greylisting module was written to collect URLs found in the body of emails as the most

common type of payload. Other payloads found in the email body, such as email addresses, images and phone numbers could be included to increase the accuracy of identifying computer zombies. Email header information was not collected as a part of this research but could also be used to increase the accuracy of identifying computer zombies.

Web page scanning

The greylisting module was written to resolve the payload URL to it's web server's IP address. Additional research may be beneficial to scan the website referenced by the payload URL for malicious content. This would assist in determining if a payload and web server were malicious in nature. This in turn could be used to block emails that contained the malicious URL and blacklist computer zombies that referenced the malicious website or web server.

7. Conclusion

Greylisting can not only be used to reduce the amount of unwanted email as a filtering technique itself, it can also be used to derive more information about the sending host to assist in differentiating it from a legitimate MTA. It is clear from the analysis in this paper that a computer zombie's purpose differs significantly from a legitimate MTA's purpose as they

employ very different methods of sending email.

From the analysis of the sending attempts of thousands of zombie computers it is possible to profile computer zombies into types. Seven computer zombie types were identified and methods for applying these profiling techniques to filter out unwanted emails were investigated. Some techniques have more demonstrable effectiveness while others may only be useful if used in conjunction with other techniques. Specifically, identifying computer zombies by multiple FROM addresses is very effective (94% success) and coupled with blacklisting known bad web servers (used by 90% of hosts), could produce a combination of filters that are even more effective.

With the many techniques currently available for blocking unwanted email, the techniques discusses in this paper can add to the arsenal that already exist to further reduce the amount of unwanted email reaching the end user. Ideally this will result in less unwanted email, leading to fewer computers becoming computer zombies, therefore reducing the botnet operator's financial gain.

8. References

Anti-Abuse Working Group, (April 2008). Email Metrics Program: The Network Operators' Perspective, Report #7 – Third and Fourth Quarters 2007. Retrieved October 21st, 2008 from http://www.maawg.org/about/MAAWG_2007-Q3-4_Metrics_Report.pdf.

Marshal Threat Research & Engineering Team, (July 2008). Marshal TRACE Report, Marshal Security Threats:Email and Web Threats. Retrieved October 21st, 2008 from http://www.marshal.com/newsimages/trace/Marshal_Trace_Report-July_2008.pdf.

Computer Zombie. (2008, October 12). In Wikipedia, the free encyclopedia. Retrieved October 21, 2008 from http://en.wikipedia.org/wiki/Zombie_computer.

Greylisting. (2008, October 6). In Wikipedia, the free encyclopedia. Retrieved October 21, 2008 from http://en.wikipedia.org/wiki/Greylisting.

Harris, E. (2003, August 21). The Next Step in the Spam Control War: Greylisting. Retrieved November 18, 2008 from http://projects.puremagic.com/greylisting/whitepaper.html.

Venema, W. The Postfix Home Page. Retrieved January 5, 2009 from http://www.postfix.org/start.html

Klensin, J. (2001, April). RFC2821 - Simple Mail Transfer Protocol. Retrieved December 22, 2008 from http://www.faqs.org/rfcs/rfc2821.html.

Vierling, T. (2004, April). THE SENDMAIL MILTER PROTOCOL, VERSION 2. Retrieved January 5, 2009 from http://search.cpan.org/src/AVAR/Sendmail-PMilter-0.96/doc/milter-protocol.txt

Vierling, T. (2007, July). Ævar Arnfjörð Bjarmason / Sendmail-PMilter. Retrieved January 5, 2009 from http://search.cpan.org/~avar/Sendmail-PMilter/lib/Sendmail/PMilter.pm

Reverse DNS lookup. (2008, Decemeber 21). In Wikipedia, the free encyclopedia. Retrieved December 22, 2008 from http://en.wikipedia.org/wiki/Reverse_DNS_lookup.

Shadowserver Foundation (2007, November). Shadowserver Foundation - Information - Botnets. Retrieved December 22, 2008 from

http://www.shadowserver.org/wiki/pmwiki.php?n=Information.Botnets#botnetuses

Last Updated: November 20th, 2017

Upcoming SANS Training

Click Here for a full list of all Upcoming SANS Events by Location

SANS San Francisco Winter 2017	San Francisco, CAUS	Nov 27, 2017 - Dec 02, 2017	Live Event
SIEM & Tactical Analytics Summit & Training	Scottsdale, AZUS	Nov 28, 2017 - Dec 05, 2017	Live Event
SANS Khobar 2017	Khobar, SA	Dec 02, 2017 - Dec 07, 2017	Live Event
SANS Munich December 2017	Munich, DE	Dec 04, 2017 - Dec 09, 2017	Live Event
European Security Awareness Summit & Training 2017	London, GB	Dec 04, 2017 - Dec 07, 2017	Live Event
SANS Austin Winter 2017	Austin, TXUS	Dec 04, 2017 - Dec 09, 2017	Live Event
SANS Frankfurt 2017	Frankfurt, DE	Dec 11, 2017 - Dec 16, 2017	Live Event
SANS Bangalore 2017	Bangalore, IN	Dec 11, 2017 - Dec 16, 2017	Live Event
SANS Cyber Defense Initiative 2017	Washington, DCUS	Dec 12, 2017 - Dec 19, 2017	Live Event
SANS Security East 2018	New Orleans, LAUS	Jan 08, 2018 - Jan 13, 2018	Live Event
SANS SEC460: Enterprise Threat Beta	San Diego, CAUS	Jan 08, 2018 - Jan 13, 2018	Live Event
SEC599: Defeat Advanced Adversaries	San Francisco, CAUS	Jan 15, 2018 - Jan 20, 2018	Live Event
SANS Amsterdam January 2018	Amsterdam, NL	Jan 15, 2018 - Jan 20, 2018	Live Event
Northern VA Winter - Reston 2018	Reston, VAUS	Jan 15, 2018 - Jan 20, 2018	Live Event
SANS Dubai 2018	Dubai, AE	Jan 27, 2018 - Feb 01, 2018	Live Event
SANS Las Vegas 2018	Las Vegas, NVUS	Jan 28, 2018 - Feb 02, 2018	Live Event
Cyber Threat Intelligence Summit & Training 2018	Bethesda, MDUS	Jan 29, 2018 - Feb 05, 2018	Live Event
SANS Miami 2018	Miami, FLUS	Jan 29, 2018 - Feb 03, 2018	Live Event
SANS London February 2018	London, GB	Feb 05, 2018 - Feb 10, 2018	Live Event
SANS Scottsdale 2018	Scottsdale, AZUS	Feb 05, 2018 - Feb 10, 2018	Live Event
SANS Secure India 2018	Bangalore, IN	Feb 12, 2018 - Feb 17, 2018	Live Event
SANS Southern California- Anaheim 2018	Anaheim, CAUS	Feb 12, 2018 - Feb 17, 2018	Live Event
SANS Dallas 2018	Dallas, TXUS	Feb 19, 2018 - Feb 24, 2018	Live Event
SANS Secure Japan 2018	Tokyo, JP	Feb 19, 2018 - Mar 03, 2018	Live Event
Cloud Security Summit & Training 2018	San Diego, CAUS	Feb 19, 2018 - Feb 26, 2018	Live Event
SANS Brussels February 2018	Brussels, BE	Feb 19, 2018 - Feb 24, 2018	Live Event
SANS London November 2017	OnlineGB	Nov 27, 2017 - Dec 02, 2017	Live Event
SANS OnDemand	Books & MP3s OnlyUS	Anytime	Self Paced