NASK

CERT POLSKA

Report

Takeover of Virut domains

February 25, 2013

Contents

1 Executive summary

At the end of January and the beginning of February 2013 NASK (Research and Academic Computer Network) – the .pl ccTLD Registry – and its security team CERT Polska took over 43 `.pl` domains used to control the Virut botnet and to spread malicious applications. These actions were preceded by a detailed legal and technical analyses and were supported by Spamhaus and VirusTotal. Some of these domains, even outside `.pl` domain, were an important part of the botnet infrastructure. As a result of these actions, all traffic from infected computers to the Command and Control servers were redirected to the sinkhole server controlled by CERT Polska. The action cripples criminals ability to control infected machines and allows to gather information about infected machines. This data is shared with all interested partners. From the gathered data, on average 270 thousand unique IP addresses connect to the botnet server every day, which is a good estimation of the botnet size at the day of takeover. Almost a half of infected machines are located in three countries: Egypt, Pakistan and India. Poland is located at the 19th place on the infection scale. This report presents the actions taken by NASK, methods used to gather data and their analysis, which offer additional insight into Virut activity, including a connection to the sale of fake antivirus applications.

2 Introduction

2.1 What is Virut?

Virut is a malicious software used to control a computer without the user knowledge. Upon startup, Virut connects to an IRC server, which is controlled by the attacker. This server, in private messages, sends commands to download and run executable files from specified URLs. Using this mechanism, one can run any command and install any software on the infected machine. Virut was used, among other things, to attach advertisements to the content displayed by user. The network of infected machines (that together composed a botnet) was also used to send spam or to perform DDoS attack.

IRC servers used to manage the botnet were usually present on the `.pl` domain (e.g. `ircgalaxy.pl` or `zief.pl`), but also on the `.ru` and `.at` domains. As a fallback mechanism each Virut sample had a short list of domains on which C&C servers were located. Latest version of Virut bot also included a Domain Generation Algorithm, which allows them to redirect traffic, in case of any C&C failure, to a specially registered `.com` domain.

Virut spreads by infecting files present on the victims system. This means that a risky behaviour such as using a pendrive, network share or downloading files from an unverified sources like torrents or crack sites can lead to an infection. Newer Virut version also modify HTML files present on the victim's computer, adding code that enables a *drive-by-download* attack, i.e. installs malware automatically on a computer through a security bug in the browser or its plugin once a user visits a site. Virut was also bundled with other malicious software, which allowed for it to spread like a computer worm. Most

common case was an attack on the RPC service, after which Virut code was downloaded and executed.

2.2 Takeover of Virut domains

Actions leading to the Virut takedown described in this document were initiated by NASK. Their main goal was to remove `.pl` domains that were used to control Virut botnet and spread malicious software. These actions were preceded by detailed legal and technical analysis. First action was to gather evidence against the suspicious domains, which were undergoing changes. This data gathering process was supported by a number of partners, in particular Spamhaus and VirusTotal.

Based on the botnet analysis, CERT Polska prepared a *sinkhole* – server that emulates Command and Control behaviour. The plan was to sinkhole all of the `.pl` domains by switching their nameservers to the ones controlled by CERT Polska. This allowed to redirect all of the botnet traffic from the infected machines and remove the ability of bots to connect to the real C&C infrastructure, which was still in the hands of cybercriminals. On the evening of Janury 17th, 2013 NASK transferred 23 domains and redirected them to the sinkhole.

The second phase of the operation done together with Home.pl (Registrar used by the miscreants for many Virut domains), which involved another 15 domains, was carried out on the 18th of January with the domains pointed to the sinkhole. The domain transfer was finalized to NASK on the 21st. Spamhaus informed CERT teams in Austria and Russia that some of domains in their countries were used to control Virut botnet. All actions were finished on 6th of February, when last 5 domains where transferred to NASK from yet another partner – Consulting Service. In total, 43 domains were redirected to the sinkhole. Main factor that allowed for this operation to take place, was the policy change in the `.pl` registry regarding the domains that are used to control and spread the malicious software.

3 Sinkhole

In the following section we present information about the `sinkhole.cert.pl` server (IP address of which is `148.81.111.111`). This server receives all of the redirected traffic from the domains that were taken over. Information received from the bots was gathered and processed by CERT Polska.

3.1 Communication

Computers infected with Virut connect to the sinkhole server using domains in three ccTLDs: `.pl`, `.at` and `.ru`. Traffic connected with C&C activity was observed only on two TCP ports: 80 and 65520. This communication is sometimes encrypted with a simple stream cipher and sometimes not.

3.2 Encrypted connections

Cipher used in the Virut botnet is similar to the Vernam cipher. Figure 1 presents a schema for this cipher. We assume that $C\,[1\ldots n]$ is a ciphertext, $P\,[1\ldots n]$ is a plaintext and $K\,[1\ldots n]$ is a keystream.

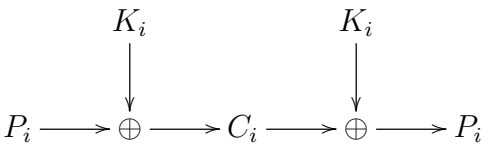

Figure 1: Vernam cipher

Vernam cipher assumes that both sides of communication have exchanged a sufficient portion of the keystream. Virut bots use a pseudorandom algorithm to generate a keystream, starting from a 4-byte number $S = (S_4S_3S_2S_1)$ generated randomly by the bot. Figure 2 illustrates the algorithm used to generate the keystream. All variables are unsigned 4 byte numbers.

$$K_i \leftarrow S_1$$
$$K_{i+1} \leftarrow S_2$$
$$S \leftarrow (S_2S_1S_4S_3)$$
$$S \leftarrow 13 \cdot S$$

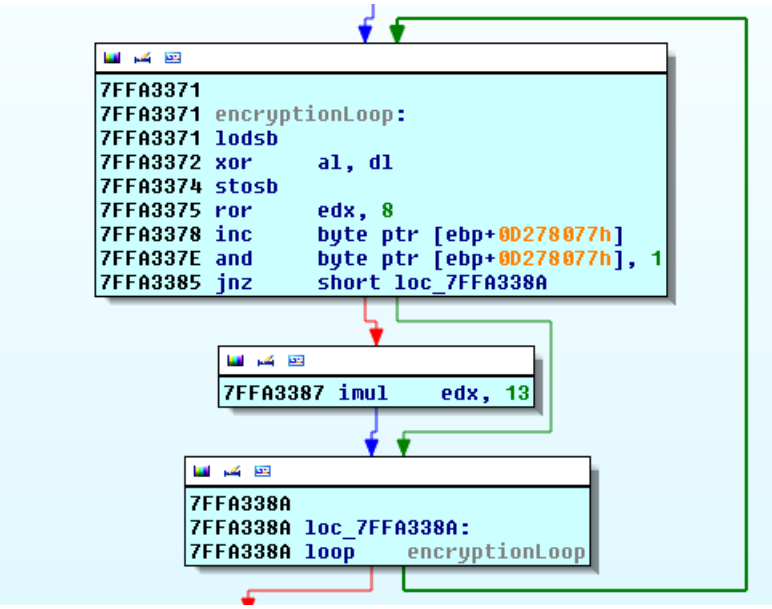

Figure 2: Keystream generation loop

First four bytes of the plaintext send by the bot are always `NICK`. Taking this into account and because first four bytes of keystream are a linear transformation of S, we can use a known plaintext attack to get S. Bots never exchange (randomly generated) S with the C&C, which leads to the conclusion that the real C&C also use known plaintext attack in order to communicate with clients.

Server responses are encrypted using the same algorithm, with the same initial value S. Backward compatibility of the server is guaranteed by the fact that unencrypted

communication can be treated as encrypted with the initial value $S = 0$.

3.3 Data sent by the bots

Infected machines, depending on the Virut bot version, after connection to the C&C server, send the following information (illustrated on figure 3):

- operating system version,
- *Volume Serial Number* (along with the checksum), i.e. a number that is assigned to the partition upon format. This number is not connected in any way with the serial number assigned by the disk manufacturer. It can also be easily changed using existing tools.
- Service Pack version.

Figure 3: Network traffic dump

Information presented in the following sections was obtained using this data. We were also able to determine that a behaviour of a particular version of bot does not differ in certain aspects. Infected machines always join the same channel and uses the same message format. Using this knowledge we have determined that there are at least 20 different versions of this malicious software. Most popular version made connections from over 1.5 million unique IP addresses (between 18th of January and 6th of February), which accounted for over a half of all connections. Some of the versions have different geographical distribution than the whole botnet, which may be a result of an effort to differentiate bots based on their country of origin. However, the most popular version did not join any channel.

3.4 Takeover of lometr.pl name server

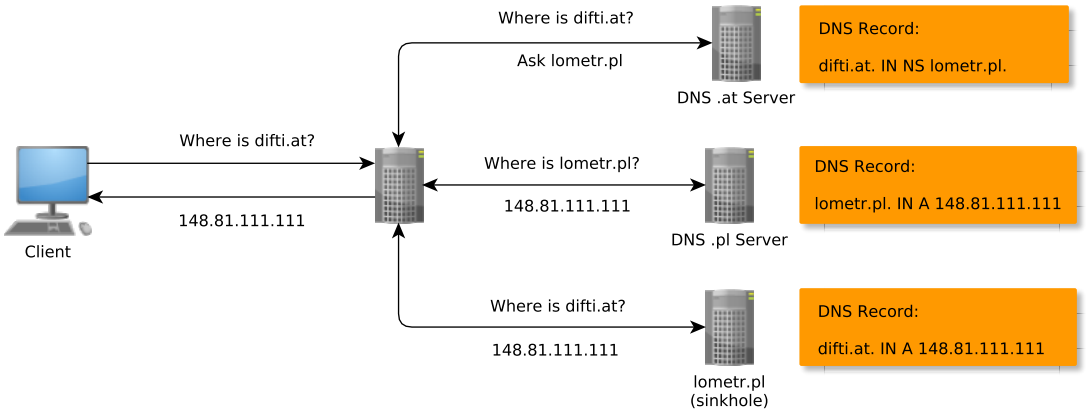

Figure 4: DNS query for .at and .ru domains

Domains that we observed as connected with the Virut botnet had `NS` Resource Record set to `*.lometr.pl` or `*.zief.pl`. This is why, when NASK took over these domains, sinkhole started to receive queries about other domains, e.g. `difti.at`. Due to this, we were able to sinkhole traffic not directed to the `.pl` domains. Figure 4 shows a DNS query for `difti.at`. Statistics presented in this report also include traffic to `.ru` and `.at` domains.

3.5 Domain Generation Algorithm (DGA)

Some versions of Virut botnet make use of the Domain Generation Algorithm (DGA) This algorithm is responsible for generating domain names, which are a backup C&Cs used when the hardcoded C&Cs have been compromised. CERT Polska analysed a Virut sample which contained the DGA.

The Virut bot uses the current date (day, month and four digit year) from the infected system. Based on this number it generates 100 6-letter domain names in the `.com` domain. Figure 5 presents a pseudocode, which is used to generate these domains. We assume that *year* is a four digit number that represents a current year, *month* and *day* represent current month and day, respectively. Result of this algorithm is saved in the, initially blank, table $domain\,[1\ldots100]$. $\mathrm{shr}\,(number, places)$ is a function that bit-shifts a *number* *places* to the right. All variables are 4 byte unsigned integers (except for the `long` casting in line 11 which is performed using 64 bit integers).

```
1:  seed ← year ∗ 10000 + month ∗ 100 + day
2:  for i ← 1 . . . 366 do                                  ▷ Creating a seed
3:      seed ← seed · 0x8088405 + 1
4:  end for
5:  for j ← 1 . . . 100 do                                  ▷ Main loop
6:      for i ← 1 . . . 594 do                              ▷ Creating a seed for domain name
7:          seed ← seed · 0x8088405 + 1
8:      end for
9:      for i ← 1 . . . 6 do                                ▷ Generating domain name
10:         seed ← seed · 0x8088405 + 1
11:         index ← shr ((long) 0x20 · seed, 32)
12:         domain[j] ← domain[j] + character [index]
13:     end for
14:     domain[j] ← domain[j] + ".com"                      ▷ Adding top level domain
15: end for
```

Figure 5: Code used to generate domain names

Character table *character* used in line 12 is defined as follows: $character\,[1 \ldots 31] = (a, \ldots, z, a, e, i, o, u, y)$. After connecting to the generated domain on port TCP/443 client expects server to authenticate itself. This is done by server sending a digitally signed message that contains a domain name that client used for connection. Pseudocode presented in figure 5 is a simplified version of the code present in the sample, fragment of which is presented on figure 6.

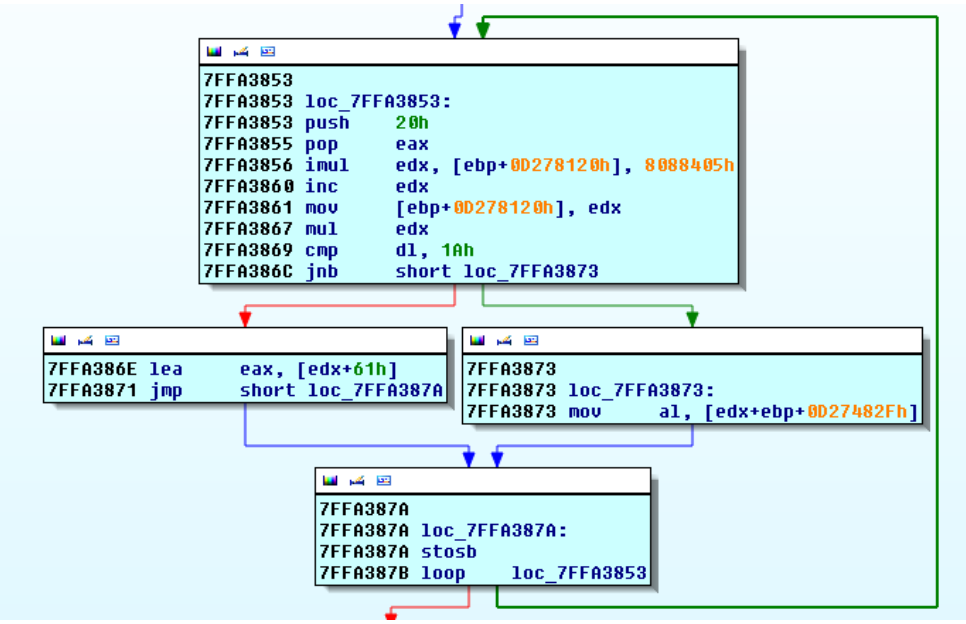

Figure 6: Main loop of the DGA algorithm

3.6 Authentication of the C&C Servers

Bot sample, analysed by CERT Polska, required authentication messages from the C&C servers. At first, C&C servers list is hardcoded into Virut sample. However, it may expand due to the Domain Generation Algorithm. In order for the generated domain to join the list, server must send a SHA-256 digest of the domain, signed using a RSA 2048 bit key, which is a part of sample code. This message must be send in a 20 second window.

When bot tries to connect to any server from the C&C list, it waits 30 seconds for the authentication procedure. C&C sends to the client current date (day, month and year) and, like in the previous case, digested and signed combination of IP server address and the aforementioned date. Client verifies the signature and then tries to match the date to the date on the infected machine. If it succeeds, authentication is finished. If the dates are different (e.g. due to the fact that machines are in different timezones) client sends a `DSTAMP` message to the server which contains a date on infected system. Bot requires C&C to sign this date (along with IP address) and send it again.

Based on the gathered data we established that not all Virut bots require this kind of authentication from the C&C.

3.7 Traffic on TCP/80 not connected to C&C

Since the beginning of the domain sinkholing we observed network traffic on port 80, that was not connected to the C&C server. Some of the domains that were taken over were used to host exploit-packs in order to infect client machines that were sending HTTP requests to them. These requests contained HTTP headers, thanks to which we are able to provide statistics for both infected and infecting hosts. Figure 7 presents a client visiting infected site `http://www.example.com/refers_virut.html`. When the machine tries to communicate with `http://www.example.com/refers_virut.html`, it first sends HTTP request to `www.example.com` starting with `GET /refers_virut.html`. If it receives an HTML file, which refers to `http://www.brenz.pl/rc`, it sends HTTP `GET /rc` request to `www.brenz.pl` server with the `Host: www.brenz.pl` header (because this request is directed to `www.brenz.pl`) and the `Referer` header is set to the original site address: `http://www.example.com/refers_virut.html` (because this is the address which referred to this resource). This request will be logged by the sinkhole server. This reference can either require user to do some action (like clicking on the link) or not (like when website is include file with JavaScript code).

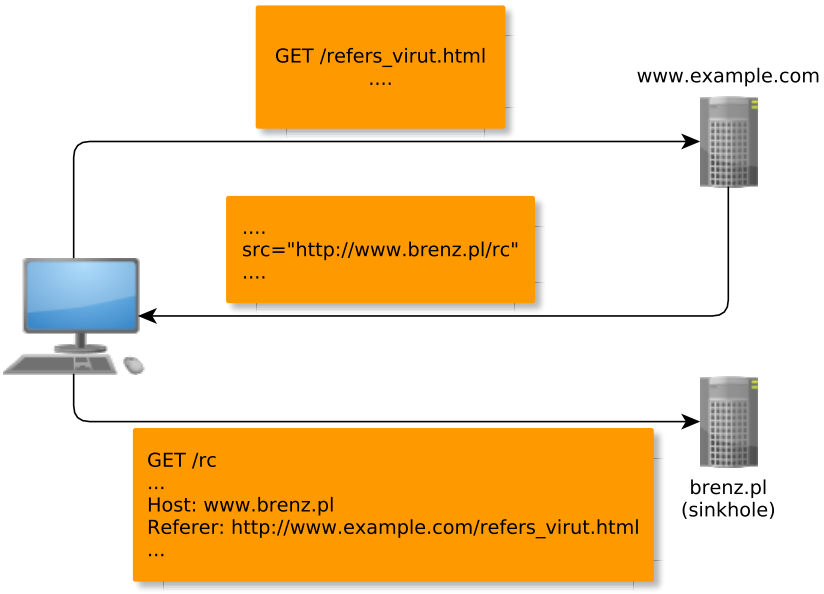

Figure 7: User connecting to http://www.example.com/refers_virut.html

Table 2 summarizes all `Host` headers send to the HTTP server. Table 1 presents 10 most popular second level domains from the `Referer` header. Among those, `whtbk.com` was present in over 27% of cases. `Referer` header contained also popular websites like `facebook.com`, `qq.com` or `youtube.com`.

	Referer	Connections
1.	whtbk.com	36 458
2.	localhost	943
3.	qqwutai.cn	498
4.	net76.net	397
5.	carlhattley.com	352
6.	hostzi.com	350
7.	07kino.com	270
8.	facebook.com	209
9.	brenz.pl	202
10.	pytiaozao.com	201

Table 1: `Referer` HTTP header

	Host	Connections
1.	brenz.pl	117 682
2.	jl.chura.pl	50 492
3.	trenz.pl	35 639
4.	zief.pl	18 847

Table 2: `Host` HTTP header

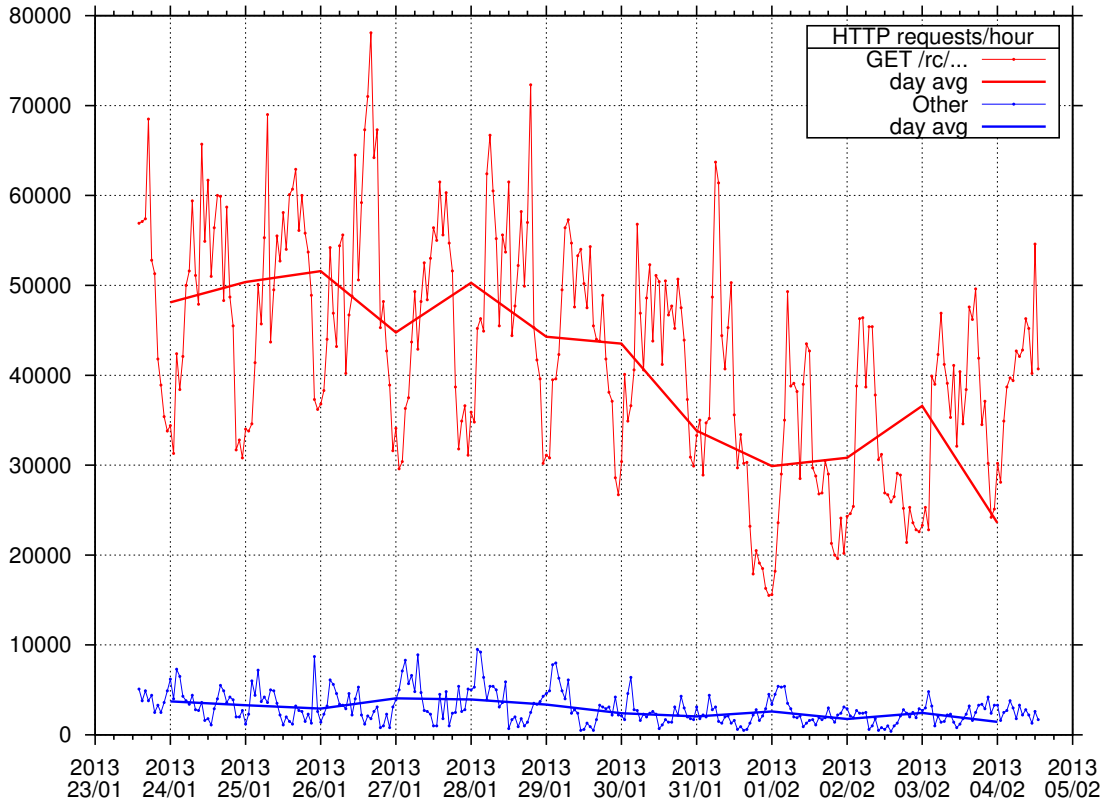

Figure 8: Number of HTTP requests over time

Besides the exploit-pack queries mentioned before, sinkhole DNS server also received queries for other websites that were connected to Virut botnet. One of these addresses was known to security researchers: `exerevenue.com`. Second one was `protsystem.com`, which contained a FakeAV. This is the kind of malicious software which poses as a trial version of the antivirus software and promises to resolve all fake infections if the user is willing to pay for an upgrade to *full version*.

CERT Polska was able to access archived version of `protsystem.com` through the `web.archive.org` service. Version presented below was available on the 4th of September 2009 and 3rd of July 2011. Figure 9 presents the main site and figure 10 presents a site set up to buy the FakeAV. 41 825 DNS `.com` queries were sent to our sinkhole between 18th of January and 14th of February and most of them involved two of the mentioned websites. Detailed breakdown of these numbers is presented in table 3.

Domain	Number of queries	Percentage
www.protsystem.com	22 665	54,19%
www.exerevenue.com	9 771	23,38%
exerevenue.com	4 990	11,93%
protsystem.com	3 945	9,43%
Other	454	1,07%

Table 3: DNS queries for `.com` domain

Figure 9: Main site of `protsystem.com`

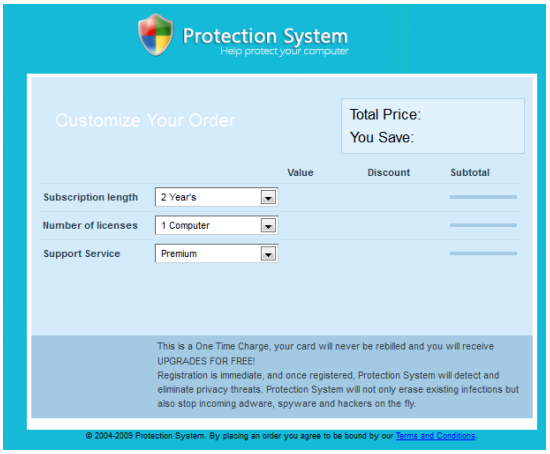

Figure 10: Site used for buying the fake antivirus

4 Statistics

In this section we present a summary of sinkhole server statistics. These include the number of connections established, geographical distribution, autonomous systems and DNS queries.

4.1 DNS Queries

DNS logs, by design, do not include IP address of the host asking for domain. Instead, they only include the IP address of the DNS resolver server. This makes it impossible to determine exactly how many individual queries where send from user machines. Number

of different IP addresses asking for domain is presented on the figure 11. On average, sinkhole registers 35 to 40 thousands of queries per day.

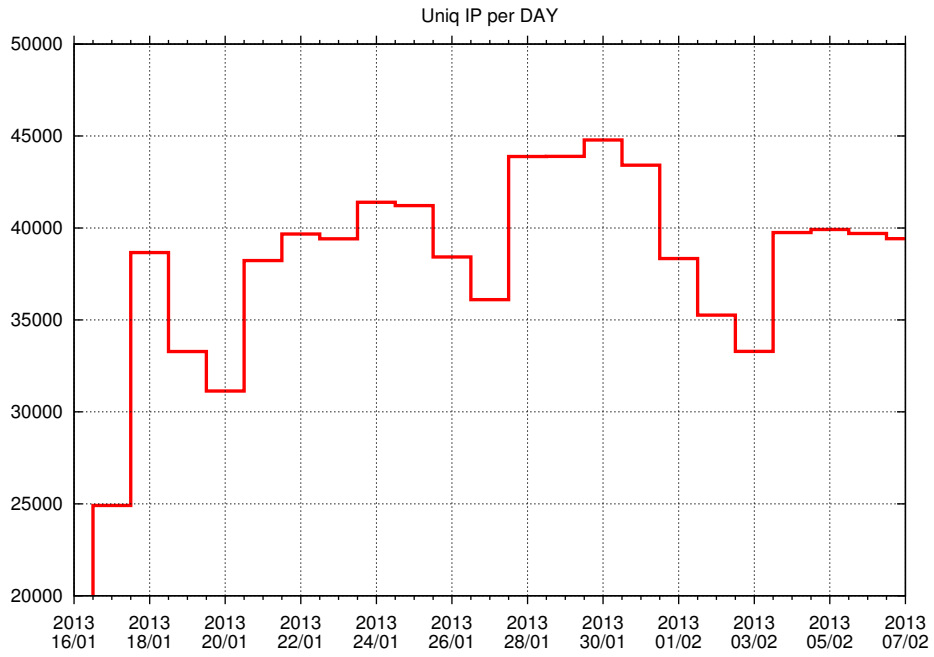

Figure 11: Number of unique IP addresses querying DNS server over time

DNS server was configured in such a way that it responded to all `A` (IPv4 query) and `NS` (nameserver) Resource Record queries. All other types of queries ended with the `NXDOMAIN` error code. As you can see on figure 12 `A` and `NS` records where queried over half of the time.

80% of all queries where connected with the `.pl` domain (figure 13). Next most queried top level domains were `.at` and `.ru`. Other top level domains appeared in logs in only 1.5‰ of cases.

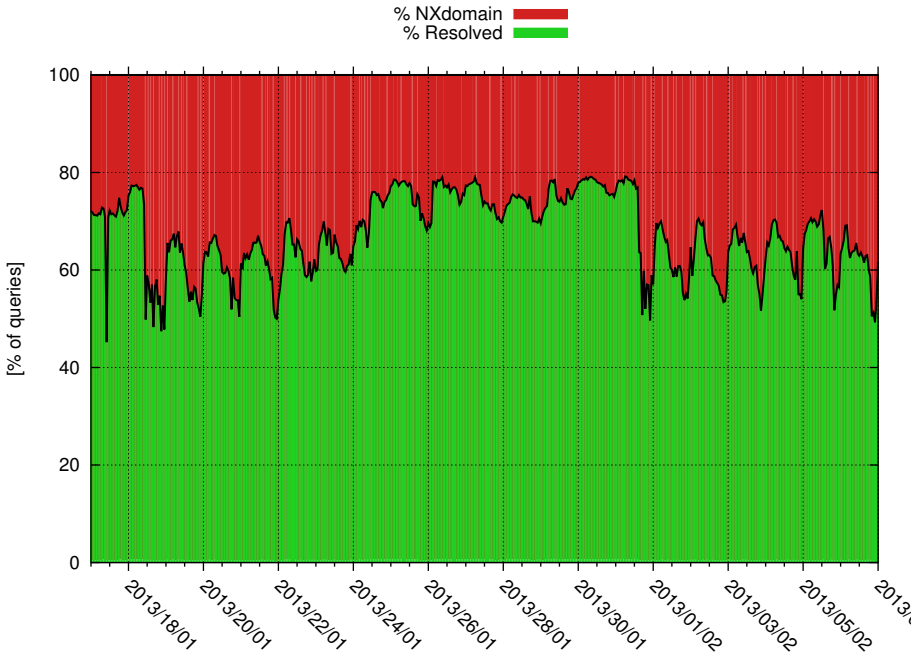

Figure 12: Responses to the DNS queries

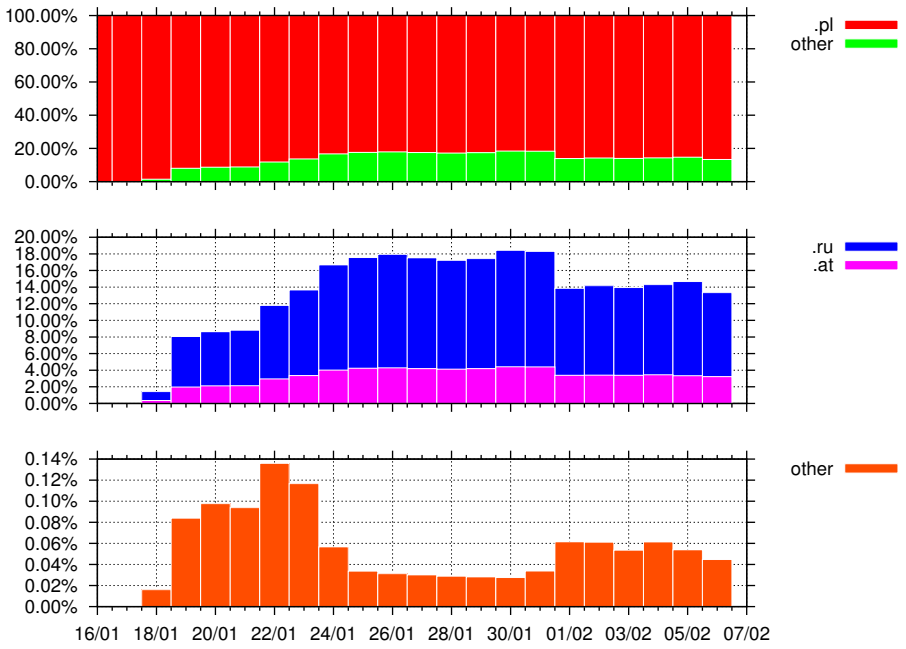

Figure 13: DNS queries divided by the top level domains

4.2 C&C connections

We have observed 3 211 135 unique IP address connecting to sinkhole from 19th of January to 5th of February. Connection were made from 218 different countries. Most of the connections were made to the TCP/80 port – from 2 657 571 unique IP addresses. Polish IP addresses was responsible for 0,67% of all connections, and 15 407 of these addresses connected to the TCP/80 port. Connection from the top 10 countries (presented in table 4) accounted for 78% of all connections. On average, 270 785 unique IP addresses connected to sinkhole every day, which allows to estimate the botnet size.

	Country	**Number of IPs**	**Percentage**
1.	Egypt	580 611	18,08%
2.	Pakistan	487 292	15,18%
3.	India	428 565	13,35%
4.	Vietnam	266 350	8,29%
5.	Iran	229 726	7,15%
6.	Indonesia	167 794	5,23%
7.	China	124 449	3,88%
8.	Algeria	92 766	2,89%
9.	Thailand	82 887	2,58%
10.	Russia	48 968	1,52%
		⋮	
19.	Poland	21 569	0,67%

Table 4: Unique IP addresses per country

Chart (figure 15) represents the number of unique IP addresses over time. Number of connections grows steadily. There were 220 598 connections to the TCP/80 port, on average, per day. Number of connections to the TCP/65520 port was substantially lower (on average 56 020 unique IP addresses per day). In total, there were 2 657 571 unique IP addresses connecting to port 80 and 748 238 to 65520.

Most of the connections were made from ISPs from Egypt and Pakistan, which is consistent with the geographical distribution of the botnet. Polish biggest ISP was on 73rd place with 5 289 infected unique IP addresses. Every day, on average, 1 647 IP addresses from Poland tried to connect to the sinkhole server. Connections were made from 268 different autonomous systems.

	Country	Number of IPs
1.	Egypt	553 593
2.	India	379 076
3.	Pakistan	284 987
4.	Vietnam	222 695
5.	Iran	212 842
6.	Indonesia	144 063
7.	China	108 520
8.	Algeria	88 906
9.	Thailand	77 578
10.	Russia	38 208
	⋮	
20.	Poland	15 407

Table 5: Connection to TCP/80

	Country	Number of IPs
1.	Pakistan	294 025
2.	Egypt	57 150
3.	India	55 132
4.	Vietnam	52 110
5.	Indonesia	34 926
6.	Iran	22 510
7.	South Africa	18 351
8.	Russia	17 648
9.	China	16 604
10.	Turkey	9 083
	⋮	
18.	Poland	6 254

Table 6: Connections to TCP/65520

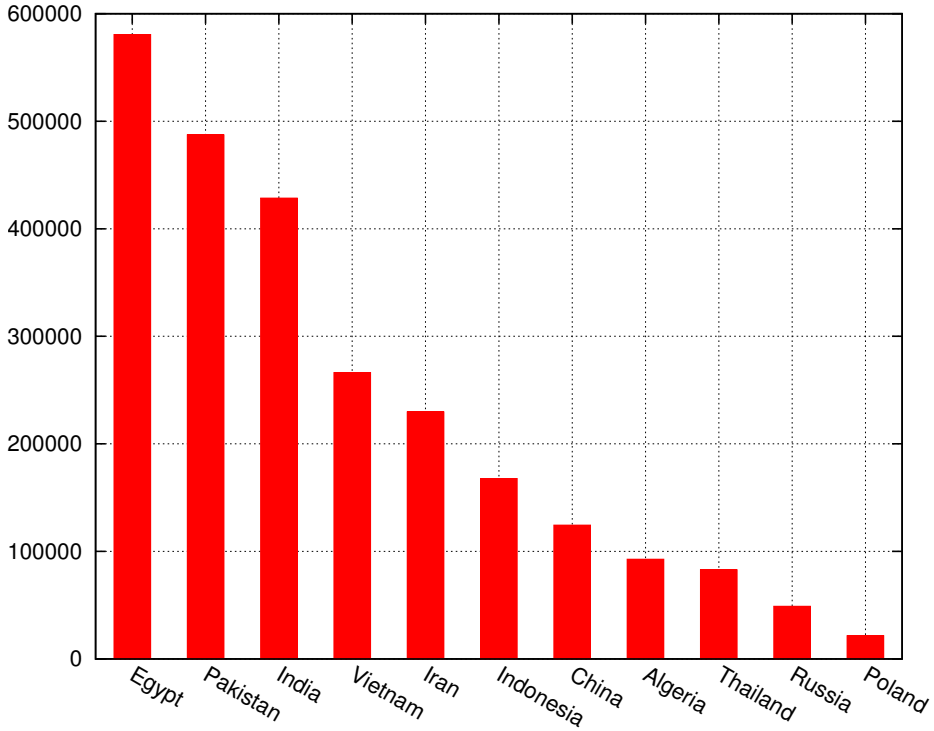

Figure 14: Geographical distribution of unique IPs in top 10 countries

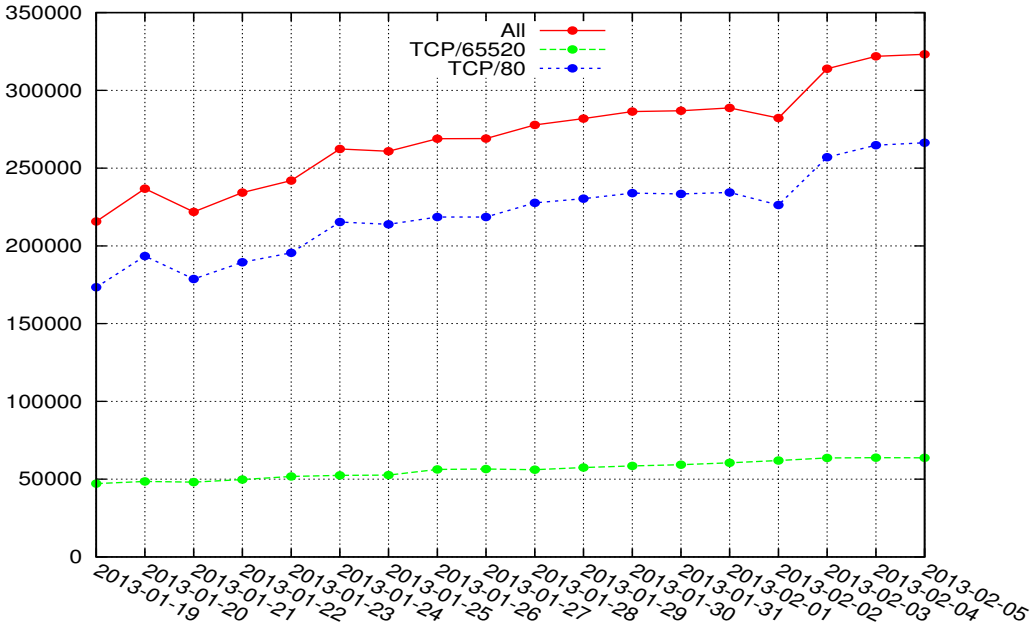

Figure 15: Number of connections from unique IPs per day

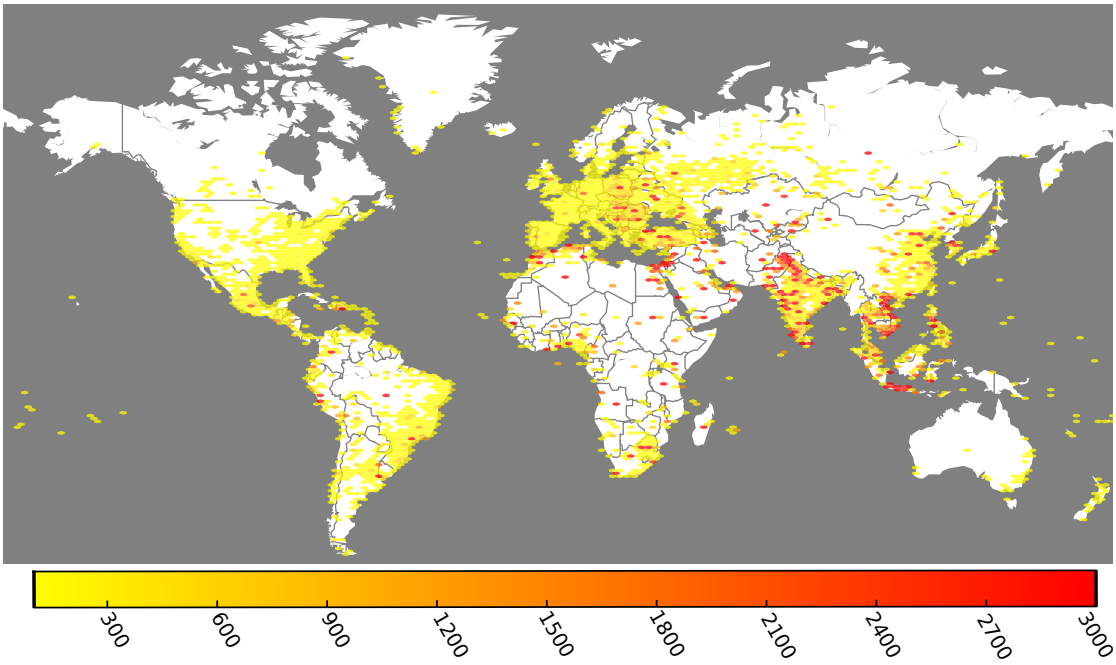

Figure 16: Geographical distribution of infected IP addresses

	Connections	ASN	AS Name	Country
1.	410 010	AS45595	Pakistan Telecom Company Limited	Pakistan
2.	364 096	AS8452	TE Data	Egypt
3.	174 002	AS45899	VNPT Corp	Vietnam
4.	145 312	AS36992	ETISALAT MISR	Egypt
5.	117 390	AS17974	PT Telekomunikasi Indonesia	Indonesia
6.	89 606	AS36947	ALGTEL-AS	Algeria
7.	79 752	AS4134	Chinanet	China
8.	75 604	AS9829	National Internet Backbone	India
9.	66 169	AS12880	Information Technology Company (ITC)	Iran
10.	65 404	AS45609	Bharti Airtel Ltd. AS for GPRS Service	India
			⋮	
73.	5 958	AS5617	Telekomunikacja Polska S.A.	Poland

Table 7: Autonomous system with the highest number of connection from unique IPs

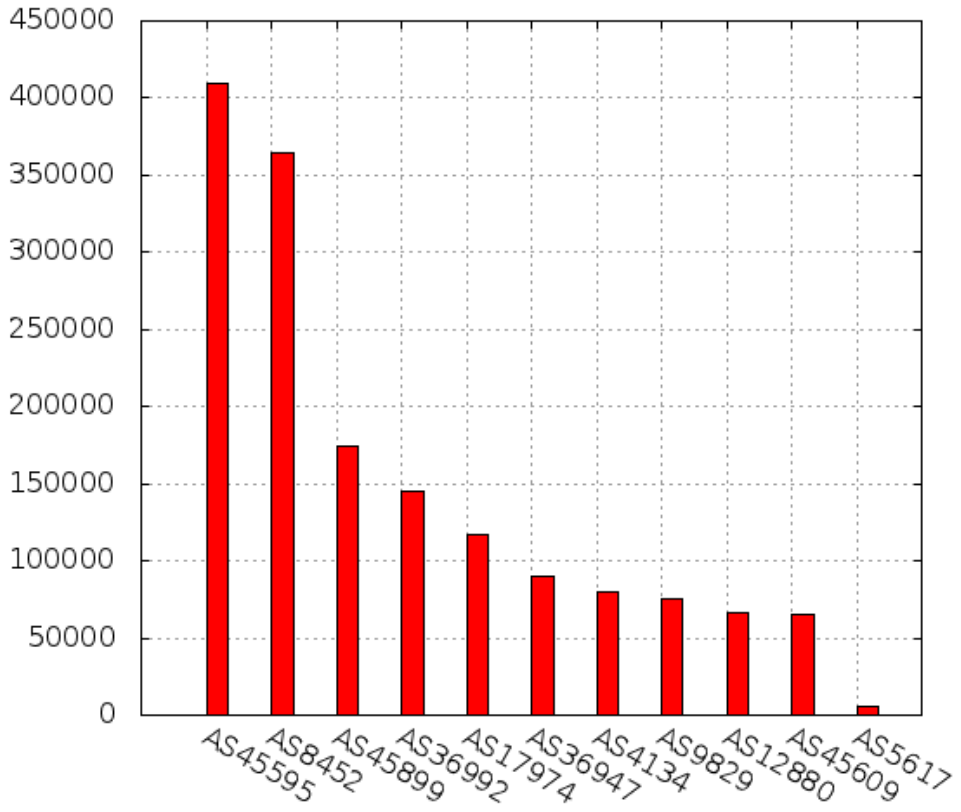

Figure 17: Autonomous systems with the highest number of connections

	Number of IPs	ASN	AS Name
1	5 958	AS5617	Telekomunikacja Polska S.A.
2	4 117	AS43447	PTK Centertel Sp. z o.o.
3	3 074	AS8374	Polkomtel Sp. z o.o.
4	2 325	AS39603	P4 Sp. z o.o.
5	1 985	AS12912	Polska Telefonia Cyfrowa S.A.
6	1 342	AS12741	Netia SA
7	1 335	AS15855	Aero 2 sp. z o.o.
8	243	AS21021	Multimedia Polska S.A.
9	132	AS6830	UPC Broadband Holding B.V.
10	89	AS29314	VECTRA S.A.

Table 8: Polish autonomous systems with the highest number of connections

Infected computers send, among other information, the version of the operating system. We observed 25 different version strings, out of which 8 indicated a valid version of Microsoft Windows. Most occurring operating system was Windows XP which accounted for over 76% connections. Second most popular was Windows 7 with over 21% of connections. All other systems were encountered only in 3% of all cases. Results presented in table 9 do not always indicate one version of the operating system. This is due to the fact that Microsoft, in order to mitigate the compatibility issues between different versions, keeps the system version string among different versions. For example, version `020601` might as well indicate Windows 7 or Windows Server 2008 R2. Some of the Virut bots version did not send any operating system version.

Operating system version	Number of occurrences
Windows XP	2 470 890
Windows 7 / Server 2008 R2	701 637
Windows Vista / Server 2008	20 792
Windows XP 64bit / Server 2003	6 038
Windows 8	3 987
Windows 2000	2 403
Windows 98	1 794
Windows ME	106
No version	28 439
Other	38

Table 9: Observed operating systems

A List of domains and IP addresses

List below presents all domains that, according to us, are connected with Virut botnet activity.

A.1 .pl domains

1. adle.pl
2. asyr.pl
3. bigex.pl
4. brans.pl
5. brenz.pl
6. bton.pl
7. cfan.pl
8. chura.pl
9. civix.pl
10. deps.pl
11. ghura.pl
12. hamb.pl
13. idet.pl
14. idon.pl
15. ircgalaxy.pl
16. ixie.pl
17. kerit.pl
18. kilme.pl
19. konter.pl
20. lifty.pl
21. lometr.pl
22. meiu.pl
23. merts.pl
24. mugu.pl
25. nels.pl
26. nigim.pl
27. play9.pl
28. ragom.pl
29. remp.pl
30. runk.pl
31. sizi.pl
32. strup.pl
33. sums.pl
34. tanz.pl
35. timid.pl
36. traum.pl
37. trenz.pl
38. tymis.pl
39. valc.pl
40. vand.pl
41. vasli.pl
42. volke.pl
43. zief.pl

A.2 .ru domains

1. alr4.ru
2. bzug.ru
3. cawt.ru
4. dbut.ru
5. gbil.ru
6. gimbs.ru
7. ijol.ru
8. ilgo.ru
9. ilopa.ru
10. ketor.ru
11. libis.ru
12. lilke.ru
13. limag.ru
14. linug.ru
15. migtu.ru
16. mlix.ru
17. pamip.ru
18. qnx1.ru
19. rdek.ru
20. rolmi.ru
21. rulm.ru
22. tasb.ru
23. tim4.ru
24. varpo.ru
25. vilq.ru
26. wict.ru
27. xalx.ru
28. xdix.ru
29. xitr.ru
30. ziten.ru

A.3 .at domains

1. amfib.at
2. difti.at
3. egab.at
4. ikepa.at
5. kamfo.at
6. maft.at
7. mampo.at
8. sox4.at

A.4 Other domains and IP addresses

Besides the domains mentioned above, there is one additional domain connected with Virut botnet: `adxhost.org`. All these domains resolved to either IP address `60.27.58.4` or one of the IP addresses from `81.177.170.0/24` class.