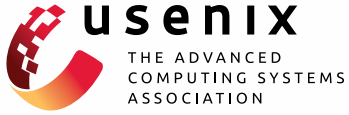


A Look at Targeted Attacks Through the Lense of an NGO

Stevens Le Blond, Adina Uritesc, and Cédric Gilbert, ***Max Planck Institute for Software Systems (MPI-SWS);*** **Zheng Leong Chua and Prateek Saxena,** ***National University of Singapore;*** **Engin Kirda,** ***Northeastern University***

https://www.usenix.org/conference/usenixsecurity14/technical-sessions/presentation/le-blond

This paper is included in the Proceedings of the 23rd USENIX Security Symposium.

August 20–22, 2014 • San Diego, CA

ISBN 978-1-931971-15-7

Open access to the Proceedings of the 23rd USENIX Security Symposium is sponsored by USENIX

A Look at Targeted Attacks Through the Lense of an NGO

Stevens Le Blond[1] Adina Uritesc[1] Cédric Gilbert[1]
Zheng Leong Chua[2] Prateek Saxena[2] Engin Kirda[3]

[1]*MPI-SWS* [2]*National Univ. of Singapore* [3]*Northeastern Univ.*

Abstract

We present an empirical analysis of targeted attacks against a human-rights Non-Governmental Organization (NGO) representing a minority living in China. In particular, we analyze the social engineering techniques, attack vectors, and malware employed in malicious emails received by two members of the NGO over a four-year period. We find that both the language and topic of the emails were highly tailored to the victims, and that sender impersonation was commonly used to lure them into opening malicious attachments. We also show that the majority of attacks employed malicious documents with recent but disclosed vulnerabilities that tend to evade common defenses. Finally, we find that the NGO received malware from different families and that over a quarter of the malware can be linked to entities that have been reported to engage in targeted attacks against political and industrial organizations, and Tibetan NGOs.

1 Introduction

In the last few years, a new class of cyber attacks has emerged that is more targeted at individuals and organizations. Unlike their opportunistic, large-scale counterparts, *targeted attacks* aim to compromise a handful of specific, high-value victims. These attacks have received substantial media attention, and have successfully compromised a wide range of targets including critical national infrastructures [19], Fortune 500 companies [23], news agencies [20], and political dissidents [10, 11, 16].

Despite the high stakes involved in these attacks, the ecosystem sustaining them remains poorly understood. The main reason for this lack of understanding is that victims rarely share the details of a high-profile compromise with the public, and they typically do not disclose what sensitive information has been lost to the attackers. According to folk wisdom, attackers carrying out targeted attacks are generally thought to be state-sponsored. Examples of national organizations that have been reported to be engaged in targeted attacks include the NSA's office of Tailored Access Operations (TAO) [3] and the People's Liberation Army's Unit 61398 [15]. Recently, researchers also attributed attacks in the Middle East to the governments of Bahrain, Syria, and the United Arab Emirates [16].

There now exists public evidence that virtually every computer system connected to the internet is susceptible to targeted attacks. The Stuxnet attack even successfully compromised air-gapped Iranian power plants [19] and was able to damage the centrifuges in the facility. More recently, Google, Facebook, the New York Times, and many other global companies have been compromised by targeted attacks. Furthermore, political dissidents and Non-Governmental Organizations (NGOs) are also being targeted [10, 11, 16].

In this paper, we analyze 1,493 suspicious emails collected over a four-year period by two members of the World Uyghur Congress (WUC), an NGO representing an ethnic group of over ten million individuals mainly living in China. WUC volunteers who suspected that they were being specifically targeted by malware shared the suspicious emails that they received with us for analysis. We find that these emails contain 1,176 malicious attachments and target 724 unique email addresses belonging to individuals affiliated with 108 different organizations. This result indicates that, despite their targeted content, these attacks were sent to several related victims (e.g., via Cc). Although the majority of these targeted organizations were NGOs, they also comprised a few high-profile targets such as the New York Times and US embassies.

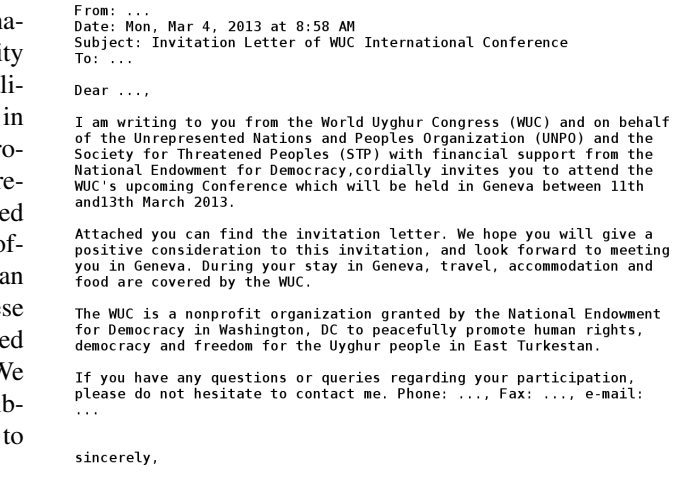
We leverage this dataset to perform an empirical analysis of targeted attacks in the wild. First, we analyze the engineering techniques and find that the language and topic of the malicious emails were tailored to the mother tongue and level of specialization of the victims. We also find that sender impersonation was common and that some attacks in our dataset originated from compromised email accounts belonging to high-profile ac-

tivists. Second, whereas recent studies report that malicious archives and executables represented the majority of the targeted-attack threat [15, 22], we find that malicious documents were the most common attack vector in our dataset. Although we do not find evidence of zero-day vulnerabilities, we observe that most attacks used recent vulnerabilities, that exploits were quickly replaced to adapt to new defense mechanisms, and that they often bypassed common defenses. Third, we perform an analysis of the first-stage malware delivered over these malicious emails and find that WUC has been targeted with different families of malware over the last year. We find that over a quarter of these malware samples exhibited similarities with those used by entities reported to have carried out targeted attacks.

Our work complements existing reports on targeted attacks such as GhostNet, Mandiant, and Symantec Internet Security Threat (ISTR) 2013 [11, 15, 22]. Whereas the GhostNet and Mandiant reports focus on the attack lifecycle *after* the initial compromise, this study provides an in-depth analysis of the reconnaissance performed *before* the compromise. We note that both approaches have pros and cons and are complementary: While it is hard for the authors of these reports to know *how* a system became compromised in retrospect, it is equally hard for us to know *if* the observed attacks will compromise the targeted system(s). Finally, whereas ISTR provides some numbers about reconnaissance analysis for industrial-espionage attacks [22], we present a thorough and rigorous analysis of the attacks in our dataset.

Finally, to foster research in this area, we release our dataset of targeted malware to the community [4].

Scope. Measuring real-world targeted attacks is challenging and this paper has a number of important biases. First, our dataset contains mainly attacks against the Uyghur and human-rights communities. While the specifics of the social engineering techniques (e.g., use of Uyghur language) will vary from one targeted community to another, we argue that identifying commonly used techniques (e.g., topic, language, senders' impersonation) and their purpose is a necessary step towards designing effective defenses. Another limitation of our dataset is that it captures only targeted attacks carried out over email channels and that were detected by our volunteers. Although malicious emails seem to constitute the majority of targeted attacks, different attack vectors such as targeted drive-by downloads are equally important. Finally, we reiterate that the goal of this study is to understand the reconnaissance phase occurring *before* a compromise. Analyzing second-stage malware, monitoring compromised systems, and determining the purpose of targeted attacks are all outside of the scope of this paper and are the topic of recent related work [10, 16]. We discuss open research challenges in Section 6.



```
From: ...
Date: Mon, Mar 4, 2013 at 8:58 AM
Subject: Invitation Letter of WUC International Conference
To: ...

Dear ...,

I am writing to you from the World Uyghur Congress (WUC) and on behalf
of the Unrepresented Nations and Peoples Organization (UNPO) and the
Society for Threatened Peoples (STP) with financial support from the
National Endowment for Democracy,cordially invites you to attend the
WUC's upcoming Conference which will be held in Geneva between 11th
and13th March 2013.

Attached you can find the invitation letter. We hope you will give a
positive consideration to this invitation, and look forward to meeting
you in Geneva. During your stay in Geneva, travel, accommodation and
food are covered by the WUC.

The WUC is a nonprofit organization granted by the National Endowment
for Democracy in Washington, DC to peacefully promote human rights,
democracy and freedom for the Uyghur people in East Turkestan.

If you have any questions or queries regarding your participation,
please do not hesitate to contact me. Phone: ..., Fax: ..., e-mail:
...


sincerely,
```

Figure 1: Screenshot of a malicious email with an impersonated sender, and a malicious document exploiting Common Vulnerabilities and Exposures (CVE) number 2012-0158 and containing malware. **The email replays an actual announcement about a conference in Geneva and was edited by the attacker to add that all fees would be covered.**

2 Overview

Context. WUC, the NGO from which we have received our dataset, represents the Uyghurs, an ethnic minority concentrated in the Xinjiang region in China. Xinjiang is the largest Chinese administrative division, has abundant natural resources such as oil, and is China's largest natural gas-producing region. WUC frequently engages in advocacy and meeting with politicians and diplomats at the EU and UN, as well as collaborating with a variety of NGOs. Rebiya Kadeer, WUC's current president, was the fifth richest person in China before her imprisonment for dissent in 1996, and is now in exile in the US. Finally, WUC is partly funded by the National Endowment for Democracy (NED), a US NGO itself funded by the US Congress to promote democracy. (We will see below that NED has been targeted with the same malware as WUC.)

WUC has been a regular target of Distributed Denial of Service (DDoS) attacks and telephone disruptions, as well as targeted attacks. For example, the WUC's website became inaccessible from June 28 to July 10, 2011 due to such a DDoS attack. Concurrently to this attack, the professional and private phone lines of WUC employees were flooded with incoming calls, and the WUC's contact email address received 15,000 spam emails in one week.

Data acquisition. In addition to these intermittent threats, WUC employees constantly receive suspicious emails impersonating their colleagues and containing

malicious links and attachments. These emails consistently evade spam and malware defenses deployed by webmail providers and are often relevant to WUC's activities. In fact, our volunteers claim that the emails are often so targeted that they need to confirm their legitimacy with the impersonated sender in person. For example, Figure 1 shows the screenshot of such an email that replays the actual announcement for a conference in Geneva organized by WUC. As a result, WUC members are wary of any emails containing links or attachments, and some of them save these emails for future inspection. We came in contact with two WUC employees who shared the suspicious emails that they had received (with consent from WUC). The authors of this work were not involved in the data collection.

Characteristics of the dataset. The two volunteers shared with us the headers and content of 1,493 suspicious emails that they received over a four-year period. 1,178 (79%) of these emails were sent to the private email addresses of the two NGO employees from whom we obtained the data, 16 via the public email address of the WUC, and the remaining 299 emails were forwarded to them (126 of these by colleagues at WUC). Overall, 89% of these emails were received directly by our volunteers or their colleagues at WUC. As we will see below, they also contain numerous email addresses in the To and Cc fields belonging to individuals that are not affiliated with WUC.

The emails contained 209 links and 1,649 attachments, including 1,176 with malware (247 RAR, 49 ZIP, 144 PDF, and 655 Microsoft Office files, and 81 files in other formats). Our analysis revealed 1,116 *malicious emails* containing malware attachments. (We were not able to verify the maliciousness of the links as most of them were invalid by the time we obtained the data.) In the following, we analyze malicious emails exclusively and we refer to *malicious archives or documents* depending on whether they contained RAR or ZIP, PDF or Microsoft Office documents, respectively. Finally, the volunteers labeled the data wherever necessary, enabling us, for example, to establish that the sender of the emails was impersonated for 84% of the emails. Table 1 summarizes the main characteristics of these malicious emails.

Scope of the dataset. Analyzing the headers of the malicious emails revealed a surprisingly large number of recipients in the To or Cc fields. In particular, we observed that malicious emails had been sent to 1,250 unique email addresses and 157 organizations. A potential explanation for this behavior could be that the attacker tampered with the email headers (e.g., via a compromised SMTP server) as part of social engineering so these emails were only delivered to our volunteers, despite the additional indicated recipients. To test this hypothesis, we considered only those emails received directly by our volunteers, originating from well-known webmail domains (i.e., aol.com, gmx.de, gmx.com, gmail.com, googlemail.com, hotmail.com, outlook.com, and yahoo.com), and verified via Sender Policy Framework (SPF) *and* DomainKeys Identified Mail (DKIM). SPF and DKIM are methods commonly used to authenticate the sending server of an email message. By verifying that these malicious emails originated from well-known webmail servers, we obtain 568 malicious emails whose headers are very unlikely to have been tampered with by the attacker. By repeating our above analysis on these emails only, we obtain 724 unique email addresses and 108 organizations. Other organizations besides WUC include NED (WUC's main source of funding and itself funded by the US congress), the New York Times, and US embassies. In summary, while we obtained our dataset from two volunteers working for a single organization, it offers substantial coverage not only of one NGO, but also of those attacks against multiple NGOs in which attackers target more than one organization with the same email. We show the full list of organizations targeted in our dataset in Appendix A.

What are targeted attacks? There is no precise definition of targeted attacks. In this paper, we loosely define these attacks as *low-volume*, *socially engineered* communication which entices *specific* victims into installing malware. In the dataset we analyze here, the communication is by email, and the mechanism of exploitation is primarily using malicious archives or documents. A targeted victim, in this work, refers to specific individuals, or an organization as a whole. When necessary, we also use the term volunteer(s) to distinguish between our two collaborators and other victims.

The terms targeted attacks and Advanced Persistent Threats (or APTs) are often used interchangeably. As this paper focuses on the reconnaissance phase of targeted attacks (occurring before a compromise), we cannot measure how long attackers would have remained in control of the targeted systems (i.e., their persistency). As a result, we simply refer to these attacks as targeted attacks, and not APTs, throughout the rest of this paper. We discuss specific social engineering characteristics that make targeted attacks difficult to detect by unsuspecting average users in Section 3, the attack vectors used in our dataset in Section 4, and the malware families they install in Section 5. Finally, we will discuss open research challenges in Section 6.

Ethics. The dataset was collected prior to our contacting WUC and for the purpose of future security analysis. Furthermore, WUC approved the disclosure of all the information contained in this paper and requested that the organization's name not be anonymized.

Table 1: Summary of our dataset originating from two volunteers. *Malicious* indicates the fraction of emails containing malware, *Impersonated* the fraction of emails with an impersonated sender, *# recipients* and *# orgs* the number of unique email addresses that were listed in the To and Cc fields of the malicious emails and the corresponding number of organizations, respectively.

	Beginning - end	*Size*	*Malicious*	*Impersonated*	*# recipients*	*# orgs*
1st volunteer	Sept 2012 - Sept 2013	98 MB	154/241 (64%)	141/154 (92%)	124	25
2nd volunteer	Sept 2009 - Jul 2013	818 MB	962/1,252 (77%)	802/962 (83%)	666	102
Total	Sept 2009 - Sept 2013	916 MB	1,116/1,493 (75%)	943/1,116 (84%)	724	108

3 Analysis of social engineering

The GhostNet, Mandiant, ISTR, and other reports [11, 15, 22] mention the use of socially-engineered emails to lure their victims into installing malware, clicking on malicious links, or opening malicious documents. For example, the GhostNet report refers to one spoofed email containing a malicious DOC attachment, and the Mandiant report to one email sent from a webmail account bearing the name of the company's CEO enticing several employees to open malware contained in a ZIP archive. Concurrent work reports the use of careful social engineering against civilians and NGOs in the Middle East [16] and also Tibetan and human-rights NGOs [10]. Despite this anecdotal evidence, we are not aware of any rigorous and thorough analysis of the social engineering techniques employed in targeted attacks. In this section, we seek to answer the following questions in the context of our dataset:

- *What social traits of victims are generally exploited?* Do attackers generally impersonate a sender known to the victim and if so who do they choose to impersonate?
- *Who are the victims?* Are malicious emails sent only to specific individuals, to entire organizations, or communities of users?
- *When are users being targeted?* When do users start being targeted? Are the same users frequently being targeted and for how long? Are several users from the same organization being targeted simultaneously?

3.1 Methodology

The analysis below focuses on 1,116 malicious emails received between 2009 and 2013.

Topics and language. To attempt to understand how well the attacker knows his victims, we manually categorized the emails (coded) by topic and language. (Unless indicated otherwise, the analysis below was performed on emails that were coded by one of the author.) The topic was determined by reading the emails' titles and bodies and, in cases where emails were not written in English, we also used an online translation service. Emails whose topic was still unclear after using the translator were labeled as *Unknown*.

Targeted victims. To determine the targeted victims of these attacks, we searched the email addresses and full names of the senders and receivers for the malicious emails originating from trustworthy SMTP servers. When available, we used their public profiles available on social media websites such as Google, Facebook, and Skype to determine their professional positions and organizations. We assume we have found the social profile of a victim if one of the three following rules applies (in that order): First, if the social profile refers directly to the email address seen in the malicious email; second, if the social profile refers to an organization whose domain matches the victims' email address; or third, if we find contextual evidence that the social profile is linked to WUC, Uyghurs, or the topic of the malicious email. Out of 724 victims' email addresses, we found the profile of 32% (237), 4% (30), and 23% (167) using the first, second, and last rule, respectively.

Organizations and industries. In the following, *WUC* refers to victims directly affiliated with the organization (including our volunteers). Other *Uyghur NGOs* include Australia, Belgium, Canada, Finland, France, Japan, Netherlands, Norway, Sweden, and UK associations. *Other NGOs* include non-profit organizations such as Amnesty International, Reporters Without Borders, and Tibetan NGOs. *Academia*, *Politics*, and *Business* contain victims working in these industries. Finally, *Unknown* corresponds to victims for which we were not able to determine an affiliation.

Ranks. We also translated the professional positions of the victims into one of the three categories: *High*, *Medium*, and *Low* profile. We consider professional leadership positions such as chairpersons, presidents, and executives as high-profile, job positions such as assistants, and IT personnel as medium-profile, and unknown and shared email addresses (e.g., NGO's contact information) as low-profile.

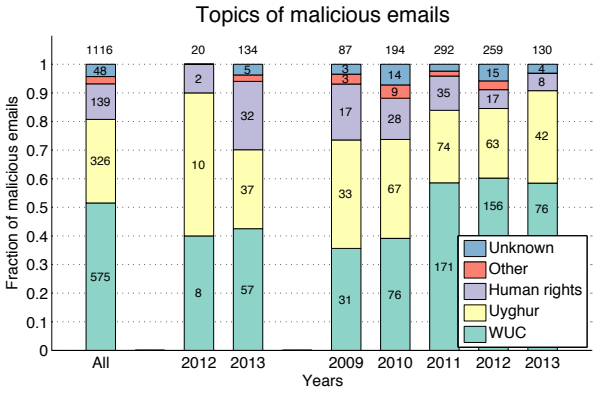


Figure 2: Distribution of the topics of the malicious emails for each year of the dataset shared by our two volunteers. The left bar corresponds to the data shared by both volunteers, and the next two bar groups to each year of the data shared by our first and second volunteer, respectively. **The content of malicious emails is targeted to the victims.**

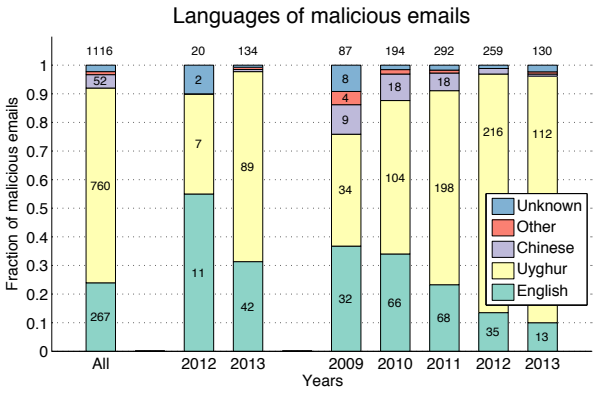


Figure 3: Distribution of languages for each year of our dataset. **Malicious emails employ the language of their victims.**

Impersonation. Finally, to understand the social context of the attack, each of our volunteers coded (based on her experience within the organization) all the email addresses of the senders into one of five categories: *Spoofed*, *Typo*, *Name*, *Suspicious*, or *Unknown*. (Coding was done based exclusively on the personal knowledge of the volunteers.) An email is marked as *Spoofed* if it bears the exact sender email address of a person known to our volunteers, as *Typo* if it resembles a sender email address known to the receiver but is not identical, and as *Name* if the attacker used the full name of a volunteer's contact (with a different email address). Finally, email addresses that look as if they had been generated by a computer program (e.g., uiow839djs93j@yahoo.com) are labeled as *Suspicious* and all remaining emails as *Unknown*. Our assumption is that, because our volunteers received most of the malicious emails directly, they were likely to recognize cases where their contacts were being impersonated. We note that labeling is conservative: Our volunteers may sometimes label *Spoofed* or *Typo* addresses as *Unknown* because they do not know the person impersonated in the attack. This may happen, for example, in cases where they were not the primary target of the attack (e.g., they appeared in Cc).

Limitations. Our dataset originates from WUC and is limited to those victims that were targeted together with that organization. We will see that these victims were often NGOs. As a result, the social engineering techniques observed here may differ from attacks against different entities such as companies, political institutions, or even other NGOs. Despite these limitations, we argue that this analysis is an important first step towards understanding the human factors exploited by targeted attacks.

3.2 Results

In this subsection, we discuss the results of our analysis of the social engineering techniques used in the malicious emails.

Topics and language. The topic of malicious emails in our dataset can generally be classified into one of three categories: WUC, Uyghur, and human-rights. In particular, we observed 51% (575) of malicious emails pertaining to WUC, 29% (326) to Uyghurs, 12% (139) to human-rights, and 3% (28) to other topics. In addition, the native language of the victim is often used in the malicious emails. In fact, 69% (664) of the emails sent to the second volunteer were written in the Uyghur language, and 62% (96) for the first one. These results indicate that attackers invested significant effort to tailor the content of the malicious emails to their victims, as we see in Figure 2 and Figure 3.

Specialized events. In addition to being on topic, we also observed that emails often referred to specific events that would only be of interest to the targeted victims. Throughout our dataset, we found 46% of events (491) related to organizational events (e.g., conferences). We note that these references are generally much more specialized than those used in typical phishing and other profit-motivated attacks. For example, Figure 1 shows a screenshot of an attack that replayed the announcement of a conference on a very specialized topic. The malicious email was edited by the attacker to add that all fees would be covered (probably to raise the target's interest).

Impersonation. We find that attackers used carefully crafted email addresses to impersonate high-profile identities that the victims may directly know. That is, attackers used one of the following four techniques to add legitimacy to a malicious email: First, 41% (465) of the

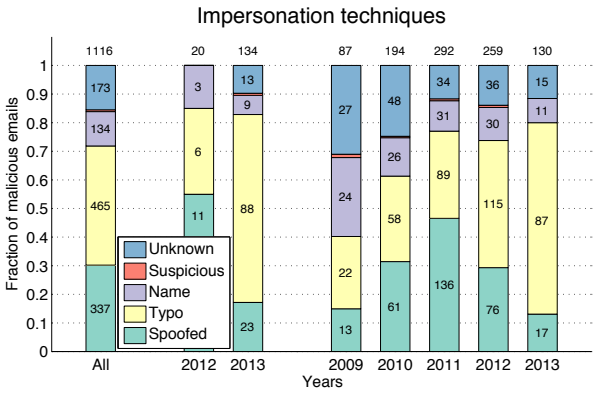


Figure 4: Distribution of senders' impersonation techniques for each year of our dataset. **Malicious emails spoof the email address of a contact of the volunteers, use a very similar address controlled by the attacker, or a contact's full name.**

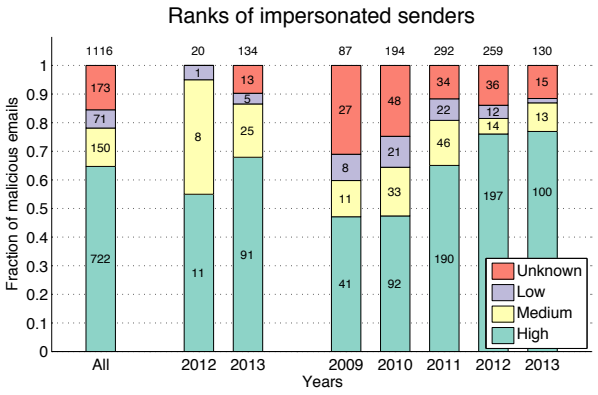


Figure 5: Distribution of impersonated senders' ranks for each year of our dataset. **Malicious emails often impersonate high-profile individuals.**

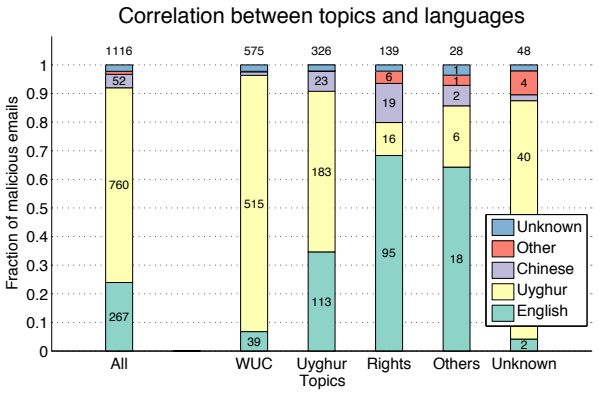


Figure 6: Distribution of languages employed to write about the main topics of malicious emails. **There is a strong correlation between malicious emails' topics and the language in which they are written.**

email addresses have *Typos* (i.e., the email address resembles known sender addresses, but with minor, subtle differences). These email addresses are identical to legitimate ones with the exception of a few characters being swapped, replaced, or added in the username. Second, 12% (134) of the senders' full names corresponded to existing contacts of the volunteers. Third, we find that most email addresses belonged to well-known email providers — Google being the most prominent with 58% of all emails using the Gmail or GoogleMail domains, followed by Yahoo with 16%.

Fourth, we find that 30% (337) of the sender emails were spoofed (i.e., the email was sent from the address of a person that the volunteer knows). This observation suggests that the attacker had knowledge of the victim's social context, and had either spoofed the email header, or compromised the corresponding email account. To identify a subset of compromised email accounts, we consider spoofed emails authenticated by the senders' domains using both SPF and DKIM. To reduce the chances of capturing compromised servers instead of compromised accounts, we also consider only well-known, trustworthy domains such as GMail. This procedure yields malicious emails that were likely sent from the legitimate account of the victims' contacts. We found that three email accounts belonging to prominent activists, including two out of 10 of the WUC leaders, were compromised and being used to send malicious emails. We have alerted these users and are currently working with them to deploy defenses and more comprehensive monitoring techniques, as we will discuss in Section 6.

We show the distributions of malicious emails sent with spoofed, typo, suspicious, or unknown email addresses in Figure 4, and the ranks of the impersonated senders in Figure 5. (We do not show the corresponding ranks for receivers because NGOs generally function with a handful of employees, all playing a key role in the organization.)

Targeted victims. For the analysis below, which leverages other recipients besides our two volunteers, we further filter emails to keep only those originating from well-known domains (as described in Section 2). Doing this leaves us with 568 malicious emails that are likely to have indeed been sent to all the email addresses in the header. We find that the attacks target more organizations than WUC, including 38 *Uyghur NGOs*, 28 *Other NGOs*, as well as 41 *Journalistic*, *Academic*, and *Political* organizations. (See Appendix A for the complete list of targeted organizations.) Interestingly, we find a strong correlation between the topic of an email and the language in which the email was written, as we show in Figure 6. Our results show that English was more and

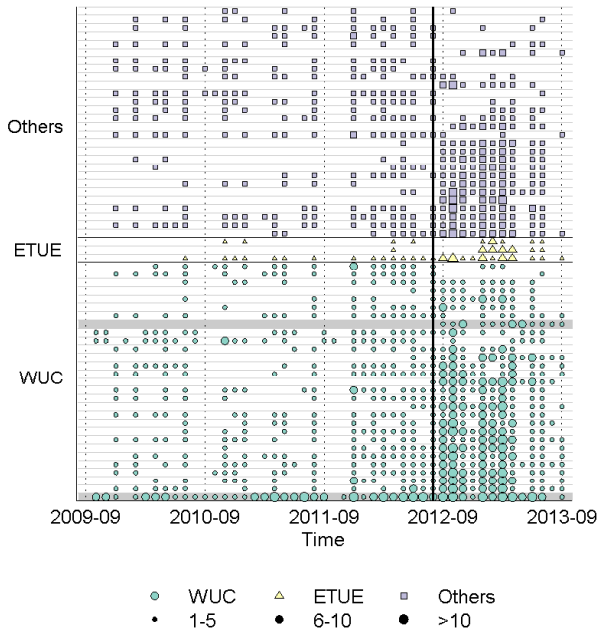


Figure 7: Timeline of attacks, in number of malicious emails per month, against the 60 most targeted victims (our two volunteers' rows are shaded and the vertical line corresponds to one of our volunteer joining WUC). The Y axis represents victims grouped by organization. *ETUE* corresponds to the East Turkestan Union in Europe NGO and *Others* to different organizations. **Each of the top 60 victims has been frequently attacked over the last four years and several victims from the same NGOs were attacked simultaneously.**

more common as the topic became less and less specialized. We hypothesize that attackers may have sent the same email messages to several recipients with similar interests to reduce the costs involved in manually crafting these emails.

Timing. Our dataset shows that the same victims were frequently targeted and that several members of the same organization were routinely targeted simultaneously. This suggests that attackers were using a "spray" strategy, trying to find the weakest links in the targeted organization, and hence, optimizing their chance of success. Spraying is clearly visible in Figure 7 where we see that the top 60 most targeted victims in our dataset received malicious emails often over the last four years. (We note that the dataset shared by one of our volunteers starts on August 2012, explaining why we observe more malicious emails after that date.) We also see that, 31 email accounts from individuals without affiliation to WUC were often targeted simultaneously to the WUC accounts.

Summary of Findings. We now revisit the initial questions posed at the beginning of this section. First, we saw that most emails in our dataset pertained to WUC, Uyghurs, or human-rights, were written in the recipient's mother tongue, and often referred to very specialized events. We also found that sender impersonation was common and that some email accounts belonging to WUC's leadership were compromised and used to spread targeted attacks. (We note that many more accounts may be compromised but remain dormant or do not appear as compromised in our dataset.) Second, we showed that numerous NGOs were being targeted simultaneously with WUC and that the specialization of emails varied depending on the recipient(s). Finally, we observed that the most targeted victims received several malicious emails every month and that attacks were sprayed over several organizations' employees.

4 Analysis of attack vectors

We now analyze the techniques used to execute arbitrary code on the victim's computer. The related work reports the use of malicious links, email attachments, and IP tracking services [10, 16]. Whereas ISTR 2013 reports that EXE are largely used in targeted attacks, and the Mandiant report that ZIP is the predominant format that they have observed in the last several years, we find that these formats represent 0% and 4% (49) of malicious attachments in our dataset, respectively. Instead, we find RAR archives and malicious documents to be the most common attack vectors. Hypotheses that may explain these discrepancies with the Mandiant report include the tuning of attack vectors to adapt to the defenses mechanisms used by different populations of email users (e.g., NGOs vs. corporations); Mandiant's attacker (APT1), mainly using primitive attack vectors such as archives; and/or Mandiant having excluded more advanced attack vectors, such as documents, from its report. However, in the absence of empirical data on APT1's attack vectors, we cannot test these hypotheses. In this section, we perform a quantitative study of the attack vectors employed in our dataset, and also analyze their dynamics. We seek to answer the following questions:

- *What attack vectors are being employed against WUC?* Do they generally rely only on human failures or also on software vulnerabilities? Do they evolve in time and if so, how quickly do they adapt to new defense mechanisms?
- *What is the efficacy of existing countermeasures?* As all malicious documents in our dataset used well-known vulnerabilities, would commercial, state-of-the-art defenses have detected all of them?

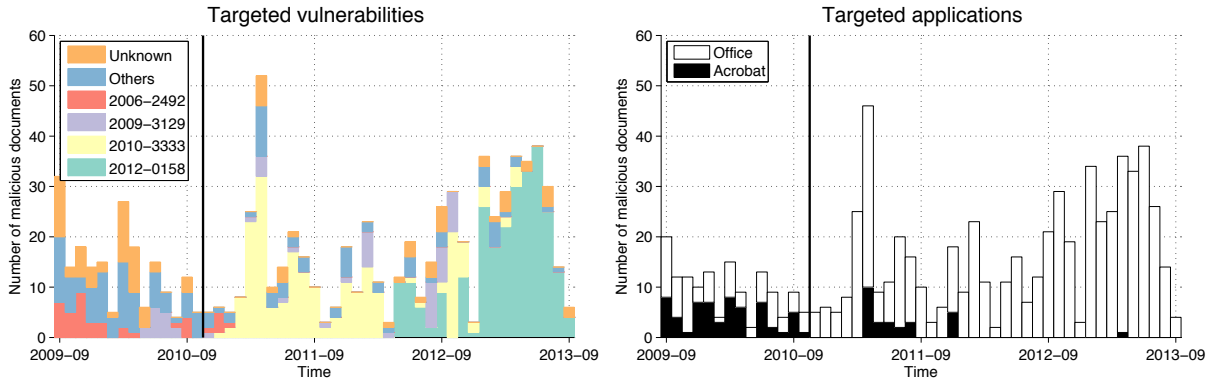


Figure 8: Number of malicious documents containing a given vulnerability (CVE) (left) and target application (right) for each month of our dataset. We represent the top four CVEs in number of attacks over the whole trace individually and *others* are represented in aggregate. The vertical line in November 2010 corresponds to the deployment of sandboxing in Acrobat Reader. **Although Acrobat Reader was the most targeted application until 2010, recent attacks mainly target the Office suite.**

4.1 Methodology

Malicious archives. To analyze the archives' contents, we extracted them in a disconnected VM environment and manually inspected their contents to determine the type of files they contain, independently of their extensions. In the case of EXE files, we also examined them manually to determine whether their Microsoft Windows icons were similar to those used for other file formats (e.g., JPEG) in order to persuade average users that they were not executable.

Malicious documents. We used two methodologies to determine the characteristics of the vulnerabilities being exploited by malicious documents. First, we submitted the documents to VirusTotal [1] for analysis. Each of the 45 Antivirus (AVs) on VirusTotal classified the checked sample as benign or malicious, and attached a "tag" describing the auxiliary information relating to the sample. Often the tag is a Common Vulnerabilities and Exposures (CVE) number, presumably corresponding to the signature that matched, but in some cases, the tag field is not a CVE; it is either tagged as "unknown" or contains a symptomatic description such as the inclusion of a suspicious OLE object. We refer to these three tags as *CVE*, *Unknown*, and *Heuristic*, respectively. Often all AVs reported a *Single* CVE but sometimes, they reported *Multiple*, conflicting CVEs. Once we collected all CVE tags, we then scraped the National Vulnerability Database [18] to obtain the release date and vulnerable applications for each of the CVEs that we found.

Second, we inspected the documents manually to confirm that they contain malware, and also used taint-assisted analysis both to verify the accuracy of the CVEs reported in AV reports and to investigate the presence of zero-day vulnerabilities.[1] The methodological details of our taint-assisted manual analysis are described in Appendix B.

Defenses. We performed a retrospective analysis of the protection offered by common defenses such as AV and webmail providers in the context of our malicious documents. For AV, we used VirusTotal to determine whether a malicious document is detected by the scanning engine of each AV, as described above. For webmail channels, we created an email account on GMail, Hotmail, and Yahoo, and used a dedicated SMTP server to send emails to that account with malicious documents attached. We considered malicious documents delivered without modifications as undetected by the webmail defenses. Otherwise, if an email or its attachment is dropped, or if the attachment's payload is modified, we considered it as detected. The analyses based on webmails and VirusTotal were performed in November 2013 and July 2014, respectively.

Limitations. As with social engineering, our analysis of attack vectors is biased towards NGOs. In addition, the above methodology is limited to the attack vectors captured in our dataset. For example, we miss attacks against the NGOs' web servers unless the corresponding malicious link appears in the suspicious emails.

Second, our taint-assisted analysis of vulnerabilities is limited to those documents for which we were able to analyze the logs manually. For example, we found that opening PDF files in our environment generated log files that were far too large (around 15GB in the median case) for manual analysis. As a result, we were able to manually confirm vulnerabilities only against Microsoft Office. However, despite this limitation, we were also able to determine which PDF documents contained malware through manual inspection.

[1] Hereafter, *zero-day vulnerabilities* refer to vulnerabilities that were not publicly disclosed at the time of the attack.

Table 2: List of well-known vulnerabilities exploited by malicious documents. *Release* corresponds to the release date of the vulnerability and *First* to its first exploitation in our data set (in number of days relative to the release date). *Resolved* corresponds to the number of Microsoft Office vulnerabilities that were mistagged in AV reports but that we were able to resolve using taint-assisted manual analysis.

CVE	*Release*	*First*	*Apps*	*# emails*	*Resolved*
2006-0022	06/13/06	1,191	Office	2	0
2006-2389	07/11/06	1,166	Office	18	16
2006-2492	05/20/06	1,125	Office	59	47
2007-5659	02/12/08	588	Acrobat	3	0
2008-0081	01/16/08	651	Office	1	0
2008-0118	03/11/08	1,010	Office	1	0
2008-4841	12/10/08	824	Office	1	0
2009-0557	06/10/09	405	Office	2	0
2009-0563	06/10/09	880	Office	31	0
2009-0927	03/19/09	180	Acrobat	11	0
2009-1862	07/23/09	68	Acrobat	3	0
2009-3129	11/11/09	188	Office	58	4
2009-4324	12/15/09	4	Acrobat	15	0
2010-0188	02/22/10	28	Acrobat	15	0
2010-1297	06/08/10	0	Acrobat	9	0
2010-2883	09/09/10	7	Acrobat	7	0
2010-3333	11/10/10	49	Office	220	0
2010-3654	10/29/10	0	Office	7	0
2011-0611	04/13/11	0	Acrobat	19	0
2011-0097	04/13/11	224	Office	3	0
2011-2462	12/07/11	2	Acrobat	5	0
2012-0158	04/10/12	37	Office	278	12
2013-0640	02/14/13	68	Acrobat	1	0

Finally, our defense analysis was performed in bulk, after the time of the attacks. As a result of the difference between the times of attack and analysis (up to four years for the first malicious documents), the detection rates reported hereafter should be treated as upper bounds. This is because the AV signatures at the time of the analysis were more up-to-date than they would have been at the time of the attack.

4.2 Results: Attack vectors

4.2.1 Malicious archives

We observed numerous targeted attacks leveraging social engineering and human failure to install malware on the victim's computer. In particular, we found 247 RAR and 49 ZIP containing malicious EXE. In 10 cases, the malicious archives were password protected with the password included in the email's body. We hypothesize that archiving was used as a rudimentary form of packer for the malware to evade detection by the distribution channels. Finally, we found that 20% of all EXEs contained in the archives used an icon that resembled a non-EXE, i.e., a DOC, JPEG, or PDF icon, in 20%, 19%, and 7% of the cases.

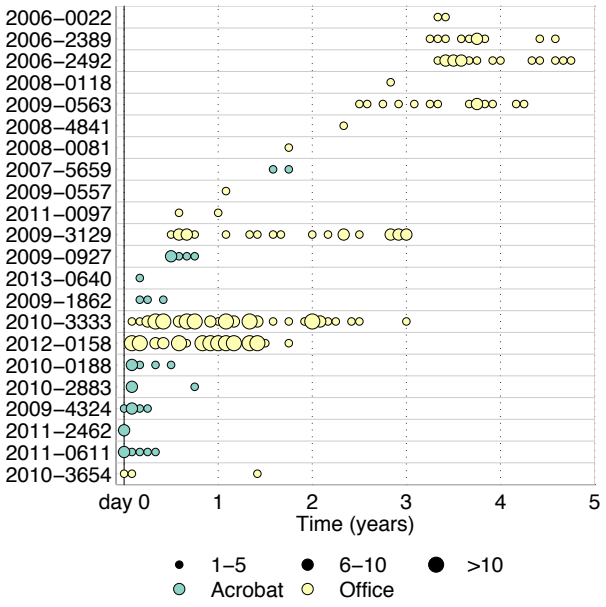


Figure 9: Timeline of the target vulnerabilities. The Y axis corresponds to CVEs and each circle to the number of CVEs seen each month after the public disclosure of the vulnerability (day 0). **All vulnerabilities were first targeted after their public disclosure.**

4.2.2 Malicious documents

We used taint-assisted analysis to resolve the conflicts due to AV mistagging and summarize the CVE information in Table 2. The number of conflicts resolved using taint-assisted manual analysis is reported in the last column *Resolved*. Additional taint-analysis results are reported in Appendix B.

Zero-day versus unpatched vulnerabilities. We find no evidence of the use of zero-day vulnerabilities against our dataset, but several uses of disclosed vulnerabilities within the same week as their public release date. In addition, we see in Figure 9 that vulnerabilities continued to be exploited for years after their disclosure, and this confirms that unpatched vulnerabilities represent a large fraction of attacks in our dataset. To ascertain the CVE being exploited in each sample, we used a combination of the telemetry data available in CVE tags generated by AVs, and a manual analysis to resolve cases where the tag was ambiguous. For each sample, we then recover the public disclosure date for the vulnerability manually, and treat it as the corresponding day-zero. By comparing the time of use in our email dataset, we are able to ascertain the lifetime of vulnerability exploits.

We find several instances of exploits that were used in publicly-reported targeted attacks in our dataset. For

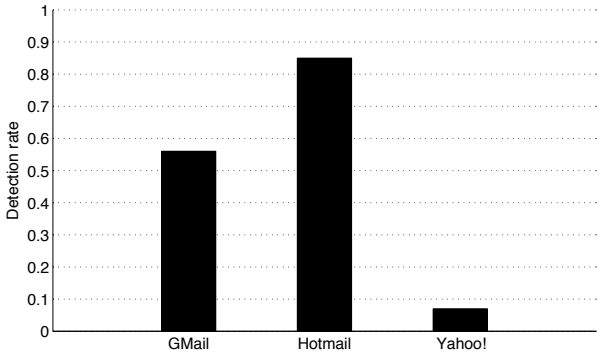


Figure 10: Detection rates of popular webmails for the malicious documents. **The efficacy of webmails to detect malicious documents varies widely.**

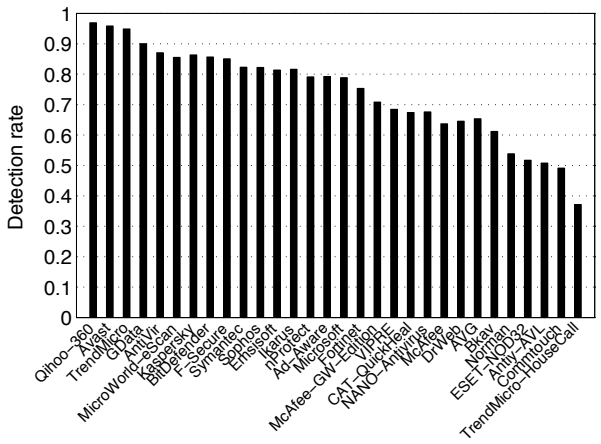


Figure 11: Detection rates of malicious documents for each of the top 30 AVs as reported by VirusTotal. **No single AV detected all malicious documents despite their use of well-known vulnerabilities.**

instance, vulnerabilities such as CVE-2009-4324, CVE-2010-3654, and CVE-2010-2883 have been reported to be zero-day vulnerabilities [6]. However, in our dataset, these vulnerabilities were used after their disclosure.

Evolution of target applications. Our data shows a sudden switch from Adobe Reader to Microsoft Office suite as the primary targeted application as of November 2010, as seen in Figure 8. We find a correlation between the time of this switch and two events: (a) the deployment of sandboxing defenses in Adobe Reader and (b) the disclosure of vulnerabilities in the Office suite. The first version of Acrobat Reader to support sandboxing for Windows (version 10.0) was released on November 15, 2010. Within the same month, a stack buffer overflow against Microsoft Office was released publicly (November 2010), reported as CVE-2010-3333. We see this CVE being massively exploited in our dataset as of January 2011, which is a time lag of two months. We observe the use of CVE-2010-3333 being replaced with CVE-2012-0158 in January 2013. This evidence suggests that attackers adapted their targeted vectors to use newly disclosed vulnerabilities within a few days to a few months of disclosure, and that updates to the security design of software reduces its exploitability in the wild (as one would expect).

4.3 Results: Bypassing common defenses

We now investigate the efficacy of existing defenses against malicious documents.

Email / Webmail Filtering. Despite the retrospective analysis of the malicious documents, we find that the detection rates of malicious documents for GMail, Hotmail, and Yahoo were still relatively low (see Figure 10). We also find that GMail failed to detect most malicious documents sent after March 2012. In particular, while the detection of documents sent before March 2012 was 73%, it is 28% after that date. Interestingly, 71% of the true positives for GMail after March 2012 corresponded to RTF files with all \r\n character sequences substituted with the \n character. While this substitution did not deactivate the malware, we observed that it broke the shellcodes embedded into these documents as they require the document size to remain unchanged to function properly. As a result, the malware was never executed. Although we cannot verify the purpose of this substitution, we note that its appearance coincided with that of the malicious RTF files. We conclude this discussion by pointing out that Yahoo's low detection rate is interesting as it claims to be using Symantec AV for its webmail service [12] — which, as we will see below, has a much higher detection rate.

Signature-based AV Scanning. In Figure 11, we show the detection rates for the top 30 VirusTotal AVs, sorted by decreasing detection rate of the malicious documents. There are two main takeaways from this graph. First, no single vendor detected all original malicious documents, even though we have seen that they used well-known vulnerabilities. For example, Qihoo, the vendor with the overall best efficacy, was unable to detect 3% of the malicious documents based on scanning. Second, we observe large variations among the efficacy of different AV vendors. That is, the detection rate dropped by 30% from the first to the twentieth AV (CAT QuickHeal) and the 15 AVs with the lowest detection rate (not shown) all had a detection rate of less than 35%.

Summary of Findings. We found that malicious documents are the most popular attack vectors in our dataset followed by malicious archives. Malicious documents tended to use newly released vulnerabilities, often within

Table 3: Summary of the malware clusters. For each cluster, we show the malware family (or an ID if we could not determine it), the number of malicious emails containing the malware, the number of Command and Control (C2) servers, the similarities in terms of communication protocols and C2 with malware attributed to known entities (*entity(Com,C2)*). **Our dataset contains several families of first-stage malware previously seen in targeted attacks carried out in the wild.**

Clusters	*1*	*2*	*Surtr*	*4*	*5*	*6*	*7*	*8*	*9*	*TravNet*
# samples	67 (9%)	58 (8%)	51 (7%)	37 (5%)	30 (4%)	22 (3%)	19 (2%)	19 (2%)	18 (2%)	13 (2%)
# C2	6	2	6	18	13	3	8	13	9	4
Similarity	DTL(C2)	—	DTL(C2)	—	—	—	—	—	—	TravNet(Com,C2)

a week, continued to utilize them for several years, and most of them used well-known instead of zero-day vulnerabilities. In particular, our taint-assisted manual analysis of Office documents did not reveal a single zero-day vulnerability in our dataset. This raises the question of whether defense mechanisms deployed in webmails and state-of-the-art commercial defenses are effective in blocking these well-known attacks. Furthermore, we found that malicious archives often contained EXE files that masquerade as pictures or documents.

5 Malware analysis

We now analyze the first-stage malware found in malicious documents. Unlike the Mandiant report, which provides an analysis for malware that targets different organizations and that (they claim) originates from the same group, our analysis focuses on all malware (in our dataset) that has targeted a single organization. By looking at targeted attacks from the perspective of the target rather than the attacker, our analysis enables us to determine whether WUC has been targeted with the same or different malware over the years. We also take a different approach from the authors of the GhostNet report who performed malware analysis on a few compromised systems belonging to different but related organizations. We instead analyze over six hundred malware samples used to establish a foothold on the targeted systems of a single organization. Our analysis differs from the related work in its scale and context [16] or focus [10]. This section aims to answer the following question:

- *Is WUC targeted with the same or different malware?* In the latter case, are there similarities between this first-stage malware and others found in targeted attacks in the wild?

5.1 Methodology

Our analysis below was done on 689 malware samples that we extracted from malicious documents.

Clustering. To make our analysis tractable for 689 malware samples, we started by clustering the malware based on its behavior. To do so, we ran the malicious EXE and DLL files in a disconnected sandboxed environment and hooked the function calls to resolve domain names and establish network communications. In addition, to obtain the TCP port number on which communication is done, we intercepted function calls to `gethostbyname` and returned a dummy routable IP address. As a result, the malware subsequently reveals the port number when it initiates a connection with the returned IP. (See Appendix C for the complete list of Command and Control (C2) domains.) Finally, we generated behavioral profiles for 586 samples, clustered them using an approach similar to [5, 14], and manually verified the accuracy of the resulting clusters.

Malware family and similarities. Similarly to Bailey et al. [5], we found that determining the malware family using AV signature scanning was unproductive. To determine whether our malware shares similarities with other known targeted malware, we relied on several reports on targeted attacks [9, 13]. We extracted the C2 domains and, when available, additional information about the malware (e.g., hashes and behavior) from these reports. Finally, we correlated the domains, IP addresses, hashes, and behavioral profiles with those from the reports in order to find similarities between the different sets of malware. We performed this analysis in February 2014.

Limitations. Our behavioral analysis was performed in a disconnected environment and as a result, it is limited to the first stage of the malware behavior. Studying the behavior of additional payload that would be downloaded after the compromise is beyond the scope of this paper and will be the subject of future work.

5.2 Results

We now analyze the malware clusters and their similarities with other targeted malware found in the wild.

Cluster sizes. We find that 57% of our malware belonged to the ten largest clusters (we show additional information about these clusters in Table 3). In total, five clusters (two in the top ten) used at least one of dtl6.mooo.com, dtl.dnsd.me, or dtl.eatuo.com as their

C2 domains, indicating some operational link between them. In fact, at the time of analysis, these three domains resolved into the same IP address and the malware in each cluster connected to different ports of the same server. Despite these apparent similarities, however, manual analysis of the behavioral logs revealed that their logic differed from one another, explaining their assignment to different clusters. Combined, these five clusters represented 24% of the malware that we analyzed.

Malware family and similarities. We found various degrees of similarities between our clusters and targeted attacks reported in the wild. First, the five clusters above had the same C2 as the DTL group reported by FireEye in November 2013 and that the malware was of the same family as one of these clusters' (*APT.9002,* not shown) [9]. In particular, we found that one of our samples in that cluster had the same MD5 hash as those described in the FireEye report and that eight had identical manifest resources. FireEye claims that this malware has been used in targeted attacks against various governmental and industrial organizations.

Second, malware in the *Surtr* cluster had the same behavioral profile as samples used against the Tibetan community in March 2012 [7]. Although the two sets of samples had different MD5 hashes, they both connected to the same C2 server (shared with *APT.9002*) on the same port number, and exhibited the same behavior to establish persistency on the victim's machine.

Third, our 13 TravNet samples exhibited similar behavior as those used against Indian targets in 2013 [2]. To do so, we obtained the samples used in India, generated their behavioral profiles, and compared them manually with the malware in our *TravNet* cluster. Although both sets connected to different C2 servers and exhibited variations in the way they searched the victims' file system, we found that they both used the same communication protocol with the C2.

Fourth, samples in another cluster communicated with the same C2 server and exhibited the same behavior as a Vidgrab sample found in a malicious document sent to a victim in Hong Kong in August 2013 [8].

Summary of findings. We found that WUC has been targeted with several malware families in the last year. We also showed that the *Surtr* and *APT.9002* clusters corresponded to malware that Citizenlab and FireEye identified as having targeted the Tibetan community, as well as other political and industrial organizations [7, 9]. Furthermore, 24% of our malware (including *Surtr*, *APT.9002* and three other clusters) had at least one C2 domain in common, which was identical to those of the Citizenlab and FireEye reports.

6 Future Work

Several directions for future work arise from this work. We briefly discuss them below.

Attack vectors and generalization. Our analysis is limited to attack vectors used against WUC. Similar studies on a wider range of targets would benefit understanding this emerging threat better. Further, our attack vectors distributed over email channels and have two main limitations. First, it is possible that our volunteers have been attacked via other channels besides email. Second, although we have seen various organizations targeted with the same malware as WUC, it is generally hard to determine with certainty which victims were the primary target of these attacks. Therefore, it is possible that other victims have been targeted with additional attack vectors when the attacks did not involve WUC. Further research is needed to overcome these limitations

Exploring different channels that attackers use for distributing malicious payloads is important. As a step towards this goal, we are currently collaborating with the Safebrowsing team at Google to investigate the emergent threat of watering-hole attacks. These attacks are conceptually very similar to drive-by download attacks with one key difference: They compromise *very specific* websites commonly visited by the targeted community (e.g., a company's website) and wait for victims to visit the website. As compared to spear phishing, watering-hole attacks offer the advantage of potentially targeting a fairly large number of victims (e.g., all employees of a large company) before raising suspicion. We conjecture that the small number of suspicious links in our dataset may be due to the small size of the targeted organizations and the public availability of their employees' email addresses.

Other attack vectors include but are not limited to packets injection to redirect victims to malicious servers (similar to those used in watering-hole attacks) and physical attacks on the victims' devices [3]. Detecting these attacks would require completely different methodologies than the one we used in this paper.

Monitoring. We have seen that a few high-profile members of the Uyghur community were compromised and that their email accounts were being used as stepping stones to carry out targeted attacks. Although it is possible that these email accounts were compromised via targeted attacks, we have not yet confirmed this hypothesis. More generally, we do not know yet what is the specific aim of these targeted attacks. Monitoring the full lifecycle of targeted attacks would require novel measurement systems, deployed at the end users, that can identify compromises without being detected.

Pinpointing the geolocation of attackers carrying out targeted attacks, or *attack attribution*, is another open

monitoring challenge. Marczak et al. were able to attribute targeted attacks to governments in the Middle East by analyzing relationships of cause and effect between compromises and real-world consequences [16]. In contrast to monitoring and attack attribution, this paper has presented a extensive, complementary analysis of the life cycle of targeted attacks *before* the compromise.

Large-scale malware analysis and clustering. We found it challenging to (a) cluster targeted malware and (b) locate similar samples. First, this malware sometimes exhibits significant similarity in its logic and different malware may also use the same Command and Control (C2) infrastructure. As a result, traditional clustering algorithms tend not to work very well. Second, we located similar samples based on a limited set of indicators such as C2, cryptographic hash, or YARA signatures, however, we feel that our current capability in that respect has a lot of room for improvement. We foresee that a search engine that can, for example, locate malware matching certain indicators out of an arbitrarily large corpus would be a useful instrument for researchers working on targeted attacks.

Our analysis of CVEs highlights that telemetry data from commercial AVs is not always reliable. Our analysis complemented with taint-analysis was largely manual and time-intensive. Analysis techniques to quickly diagnose known CVEs directly from given exploits is an open problem and perhaps one of independent interest.

Defenses. Our findings confirm that AVs may miss known CVEs, even years after their release dates. Clearly, known CVEs contribute a large part of the emerging threat of targeted attacks. Understanding why commercial AVs miss known attacks conclusively, for example to tradeoff false positives or performance for security, is an important research direction. Designing effective defenses against targeted attacks is a major research challenge which depends on our ability to understand the threat at hand. As part of future work, one could evaluate the effectiveness of novel defenses based on the findings from this paper. As a small step towards that goal, we plan to soon deploy a webmail plugin that combines metadata and stylometry analysis [17] to detect contact impersonation.

7 Conclusion

We have presented an empirical analysis of a dataset capturing four years of targeted attacks against a human-rights NGO. First, we showed that social engineering was an important component of targeted attacks with significant effort paid in crafting emails that look legitimate in terms of topics, languages, and senders. We also found that victims were targeted often, over the course of several years, and simultaneously with colleagues from the same organization. Second, we found that malicious documents with well-known vulnerabilities were the most common attack vectors in our dataset and that they tended to bypass common defenses deployed in webmails or users' computers. Finally, we provided an analysis of the targeted malware and showed that over a quarter of samples exhibited similarities with entities known to be involved in targeted attacks against a variety of industries. We hope that this paper, together with the public release of our malware dataset, will facilitate future research on targeted attacks and, ultimately, guide the deployment of effective defenses against this threat.

Acknowledgements. The authors would like to thank the anonymous reviewers, our shepherd Stuart Schechter, and Peter Druschel for their useful feedback. We would also like to acknowledge Karmina Aquino (F-Secure), Emiliano Martinez (VirusTotal), Mila Parkour (Contagio), and Nart Villeuneuve (FireEye) for their help locating similar malicious documents. Finally, we thank the Max Planck Society, the Ministry of Education of Singapore under Grant No. R-252-000-495-133, and the NSF under Grant No. CNS-1116777 for partially supporting this work.

References

[1] http://www.virustotal.com.

[2] Inside report APT attacks on indian cyber space. Tech. rep. http://g0s.org/wp-content/uploads/2013/downloads/Inside_Report_by_Infosec_Consortium.pdf.

[3] Inside TAO: Documents reveal top nsa hacking unit. http://www.spiegel.de/international/world/the-nsa-uses-powerful-toolbox-in-effort-to-spy-on-global-networks-a-940969.html.

[4] The Slingshot Project. http://slingshot.mpi-sws.org.

[5] BAILEY, M., ANDERSEN, J., MORLEYMAO, Z., AND JAHANIAN, F. Automated classification and analysis of internet malware. In *Proceedings of Recent Advances in Intrusion Detection (RAID 2007)*.

[6] BILGE, L., AND DUMITRAS, T. Before we knew it: An empirical study of zero-day attacks in the real world. In *Proceedings of the 2012 ACM conference on Computer and communications security (CCS 2012)* (2012).

[7] CITIZEN LAB. Surtr: Malware family targeting the tibetan community. Tech. rep. https://citizenlab.org/2013/08/surtr-malware-family-targeting-the-tibetan-community/.

[8] CONTAGIO. http://contagiodump.blogspot.de/2013/09/sandbox-miming-cve-2012-0158-in-mhtml.html.

[9] FIREEYE. Supply chain analysis: From quartermaster to sunshop. Tech. rep.

[10] HARDY, S., CRETE-NISHIHATA, M., KLEEMOLA, K., SENFT, A., SONNE, B., WISEMAN, G., AND GILL, P. Targeted threat index: Characterizing and quantifying politically-motivated targeted malware. In *Proceedings of the 23rd USENIX Security Symposium* (San Diego, CA).

[11] INFORMATION WARFARE MONITOR. Tracking ghostnet: Investigating a cyber espionage network. Tech. rep., 2009.

[12] JANA, S., AND SHMATIKOV, V. Abusing file processing in malware detectors for fun and profit. In *Proceedings of the 33rd IEEE Symposium on Security & Privacy* (San Francisco, CA).

[13] KASPERSKY. The nettraveler (aka TravNeT). Tech. rep. https://www.securelist.com/en/downloads/vlpdfs/kaspersky-the-net-traveler-part1-final.pdf.

[14] KRUEGEL, C., KIRDA, E., COMPARETTI, P. M., BAYER, U., AND HLAUSCHEK, C. Scalable, behavior-based malware clustering. In *Proceedings of the 16th Annual Network and Distributed System Security Symposium (NDSS 2009)* (2009).

[15] MANDIANT. APT1 exposing one of chinas cyber espionage units. Tech. rep., 2013. http://intelreport.mandiant.com/.

[16] MARCZAK, W. R., SCOTT-RAILTON, J., MARQUIS-BOIRE, M., AND PAXSON, V. When governments hack opponents: A look at actors and technology. In *Proceedings of the 23rd USENIX Security Symposium* (San Diego, CA).

[17] NARAYANAN, A., PASKOV, H., GONG, N. Z., BETHENCOURT, J., CHUL, E., SHIN, R., AND SONG, D. On the feasibility of internet-scale author identification. In *Proceedings of the 33rd conference on IEEE Symposium on Security and Privacy. IEEE* (San Francisco, CA, 2012).

[18] NATIONAL VULNERABILITY DATABASE. https://nvd.nist.gov/.

[19] RALPH LANGNER. To kill a centrifuge: A technical analysis of what stuxnets creators tried to achieve. Tech. rep., 2013.

[20] REUTERS. Journalists, media under attack from hackers: Google researchers. www.reuters.com/article/2014/03/28/us-media-cybercrime-idUSBREA2R0EU20140328.

[21] SONG, D., BRUMLEY, D., YIN, H., CABALLERO, J., JAGER, I., KANG, M. G., LIANG, Z., NEWSOME, J., POOSANKAM, P., AND SAXENA, P. Bitblaze: A new approach to computer security via binary analysis. In *Proceedings of the 4th International Conference on Information Systems Security* (Hyderabad, India, 2008).

[22] SYMANTEC. 2013 internet security threat report, volume 18. Tech. rep. http://www.symantec.com/security_response/publications/threatreport.jsp.

[23] WIRED. Google hackers targeted source code of more than 30 companies. http://www.wired.com/2010/01/google-hack-attack/.

A Targeted organizations

Organization	*# Recipients*	*# Emails*	*First-Last*
World Uyghur Congress (WUC)	53	2,366	2009-2013
East Turkestan Union in Europe (ETUE)	7	153	2010-2013
Australian Uyghur Association	3	129	2009-2013
Euro-Asia Foundation in Turkey	2	101	2010-2013
Uyghur Canadian Association	6	98	2009-2013
Germany Uyghur Women Committee	2	82	2009-2013
Radio Free Asia (RFA)	12	80	2010-2013
France Uyghur Association	5	80	2009-2013
Eastern Turkestan Australian Association (ETAA)	6	77	2009-2013
Uyghur American Association (UAA)	10	72	2010-2013
Eastern Turkestan Uyghur Association in Netherlands	4	69	2010-2013
Netherland Uyghur Union	1	60	2012-2013
United Nations for a Free Tibet (UK)	1	57	2011-2013
Eastern Turkestan Culture and Solidarity Association	13	53	2009-2013
Viktoria Uyghur Association	1	48	2010-2013
Japan Uyghur Association	2	43	2012-2013
Switzerland East Turkestan Association	2	43	2010-2013
Hacettepe University Turkey	3	41	2010-2013
Kazakhstan Academy of Poetry	2	36	2009-2013
Belgium Uyghur Association	1	35	2009-2013
Kyrgyzstan Uyghur Association	2	33	2011-2013
Uyghur Canadian Society	1	31	2009-2013
Uyghur Academy	5	25	2009-2013
Munich Uyghur Elders Meshrep	1	22	2012-2013
Republican Uyghur Cultural Center of Kazakhstan	1	22	2009-2013
Sweden Uyghur Association	4	12	2010-2013
Virginia Department of Social Services	1	11	2009-2012
Unrepresented Nations and Peoples Organization (UNPO)	5	8	2010-2013
Sociale Verzekeringsbank (SVB) NGO Netherland	1	8	2012-2013
China Democratic Party (CDP)	5	5	2009-2011
Finland Uyghur Association	1	5	2013
Jet Propulsion Laboratory, founded by NASA	1	5	2012
Pennsylvania State University US	1	5	2010-2013
Uyghur Support Group Nederland	2	5	2010-2013
Norway Uyghur Committee	1	5	2010-2013
Amnesty International	4	4	2010-2012
Association of European Border Regions (AEBR)	1	4	2010-2011
Howard University US	1	4	2012-2013
Initiatives for China	3	4	2009-2010
LSE Asia Research Center and Silk Road Dialogue	2	4	2012-2013
The Government-in-Exile of the Republic of East Turkistan	2	4	2010-2011
Uyghur Human Rights Project (UHRP)	2	4	2010
Australian Migration Options Pty Ltd	3	4	2010
Agence France-Presse	1	3	2013
National Endowment for Democracy (NED)	2	3	2010-2012
PEN International	3	3	2009-2013
Syracuse University US	1	3	2013
Worldwide Protest in Honor and Support of Uyghurs Dying for Freedom	1	3	2013
Australian Government - Department of Foreign Affairs and Trade	2	3	2010
New Tang Dynasty Television China	2	3	2010
The Epoch Times	2	3	2010
Ministry of Foreign Affairs Norway	1	2	2013
International University of Kagoshima Japan	1	2	2013
Association of Islam Religion	1	2	2013
Bilkent University Turkey	2	2	2011-2012
Embassy of Azerbaijan in Beijing	2	2	2010
Indiana University School of Law-Indianapolis LL.M.	1	2	2012
KYOCERA Document Solutions Development America	1	2	2013
New York Times	2	2	2009
Pfizer Government Research Laboratory - Clinical Pharmacology	1	2	2011-2012
Saudi Arabia - Luggage Bags and Cases Company	1	2	2013
Students for a Free Tibet	2	2	2010
Sweden Uyghur Education Union	1	2	2010-2013
Uyghur International Culture Center	1	2	2012
The Protestant Church Amsterdam	1	2	2010
Swiss Agency for Development and Cooperation (SDC) Kargyzstan	1	1	2013
American Bar Association for Attorneys in US	1	1	2010
Assistance for Work Germany Frankfurt	1	1	2010
Bishkek Human Rights Committee	1	1	2012
Central Tibetan Administration (CTA)	1	1	2010
Chinese Translation Commercial Business	1	1	2009
Circassian Cultural Center (CHKTS)	1	1	2012
Colombian National Radio	1	1	2010
Embassy of the United States in Australia	1	1	2010
Europa Haber Newspaper Turkey	1	1	2010
Europe-China Cultural Communication (ECCC)	1	1	2011
Freelance Reporter and writer Turkey	1	1	2012
Goethe University Frankfurt am Main Germany	1	1	2012
Human Rights Campaign in China	1	1	2010
International Enterprise (IE) - Singapore Government	1	1	2010
International Tibet Independence Movement	1	1	2010
Jasmine Revolution China (Pro-Democracy Protests)	1	1	2009
Socialist Party (Netherlands)	1	1	2011
Los Angeles Times	1	1	2010
Milli Gazete (National Newspaper Turkey)	1	1	2010
Norwegian Tibet Committee	1	1	2010
Photographer Turkey	1	1	2012
CNN International Hong Kong	1	1	2012
Reporters Without Borders	1	1	2012
Republican National Lawyers Association Maryland	1	1	2010
Save Tibet - International Campaign for Tibet	1	1	2010
Society for Threatened People (STPI)	1	1	2012
Southern Mongolian Human Rights	1	1	2012
Stucco Manufacturers Association US	1	1	2013
Superior School of Arts France	1	1	2012
The George Washington University	1	1	2013
TurkishNews Newspaper	1	1	2010
US Bureau of Transportation Statistics	1	1	2009
Umit Uighur Language School	1	1	2010
Union of Turkish-Islamic Cultural Associations in Europe	1	1	2012
University of Adelaide Melbourne	1	1	2010
University of Khartoum Sudan	1	1	2012
US Embassy and Consulate in Munich Germany	1	1	2011
Wei Jingsheng Foundation	1	1	2009
Xinjiang Arts Institute China	1	1	2010
Yenicag Gazetti (Newspaper Turkey)	1	1	2010
American University	1	1	2012
Islamic Jihad Union	1	1	2012

B Dynamic taint-assisted analysis of malicious documents

B.1 Methodology

We use BitBlaze [21] to perform dynamic taint-tracking analysis of the targeted applications under the malicious documents as input and configure it to report four kinds of reports: (a) when a tainted Extended Instruction Pointer (EIP) is executed, (b) when a memory fault is triggered in the target program, (c) when a new process is spawned from the target program, or (d) when the analysis "times out" (i.e., runs without interruption for over 15 minutes). To mark malicious documents as a source of taint, we tainted the network inputs and routed the malicious input file using `netcat`. Additionally, we set BitBlaze to exit tracing at the detection of null pointer exceptions, user exceptions, tainted EIPs, and process exits before the start of the trace. A tag was generated from the trace by obtaining the last instruction with tainted operands, and matching it with the list of loaded modules generated by TEMU. Our guest (analysis) system configuration used in the image consists of clean installations of Windows XP SP2 with TEMU drivers and Microsoft Office 2003.

B.2 Results

Anti-virus software typically uses static signature-matching or whitelisting techniques to analyze malware. To validate the analysis results available from commercial AV, we ran a separate semi-automated dynamic analysis of the targeted application under our malicious documents.

Out of 817 unique input documents (725 malicious and 92 legitimate), 295 timed out with our BitBlaze analysis without reporting a tainted EIP, a memory fault, or a newly spawned process.[2] Another 13 of them were incompatible with our analysis infrastructure (using a more recent DOCX format). We could not compare these cases directly to the results obtained from VirusTotal. Therefore, we focus on the remaining 509 malicious documents in the evaluation.

Efficacy of Taint EIP Detection. Taint-tracking detected tainted EIP execution in 477 out of the 509 documents. In 19 cases of the undetected 32 cases, however, a new process was spawned without it being detected by taint-tracking. We treat these as false negatives in taint-tracking. We speculate that this is likely to be due to missed direct flows, untracked indirect flows (via control dependencies, or table-lookups), or attacks using non-control-flow hijacking attacks (such as argument corruption). 13 documents did not lead to a tainted EIP execution, but instead caused a memory fault. This could be due to a difference in our test infrastructure and the victim's, or an attempt to evade analysis. In 33 of the 477 cases where tainted EIP was detected, no new spawned process was created, and the tainted EIP instruction did not correspond to any shellcode. All these cases correspond to a particular instruction triggering the tainted EIP detection in MSO.DLL, a dynamic-link library found in Microsoft Office installations. To understand this case better, we manually created blank benign documents and fed them to Microsoft Office — they too triggered tainted EIP detections. We treat these cases as false positives in taint detection, possibly because of benign dynamic generation of code. All the remaining cases (i.e., 444 out of 477) are legitimate exploits that we could confirm to execute shellcode.

[2] We believe 150 of these are due to user-interaction which we could not presently automate, and the remaining could potentially be analyzed with a faster test platform; we plan to investigate this in the future.

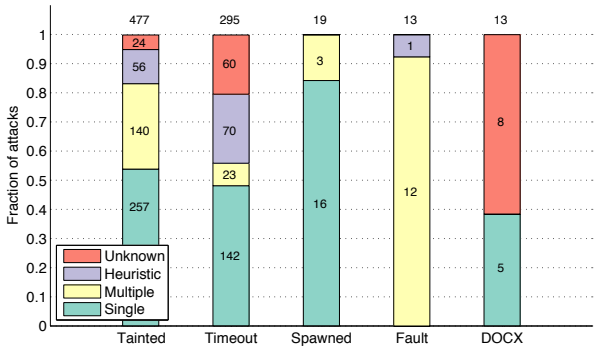


Figure 12: Breakdown of dynamic taint-assisted analysis, and comparison to VirusTotal AV results. *Single*, *Multiple*, *Heuristics*, and *Unknown* correspond to the different AV tags assigned to documents. The main bars show the detection result from BitBlaze: (a) Detected by *Tainted* EIP execution, (b) *Timeout*, (c) *Spawned* process without tainted EIP execution, (d) Memory *Fault* without tainted EIP execution, and (e) *DOCX* unable to run in our analysis environment. Within each main bar, each stacked bar represents the corresponding tag given by VirusTotal.

Dynamic Taint versus VirusTotal. Figure 12 shows the detailed comparison of taint-assisted classification of vulnerabilities versus the results from VirusTotal. Out of a total of 477 documents on which tainted EIP was detected, VirusTotal tagged 397 documents with one or more CVEs. Of the remaining 80 cases that are detected by tainted EIP execution, 24 are undetected by VirusTotal, and 56 are detected, but marked *Unknown* (i.e., no CVE assigned) by VirusTotal. Dynamic taint analysis to determine the tainted EIP was helpful to further refine the results of AV detection for a majority of these 56 tagged-*Unknown* cases. Specifically, for 55 out of the 56 documents, taint-assisted manual analysis was able to resolve it to the exploited CVE.

Out of a total of 477 documents on which tainted EIP was detected, VirusTotal tagged 397 documents with one or more CVEs. Our taint-assisted manual analysis agrees with the VirusTotal CVE tag results on 372 of these 397. That is, 372 documents were detected to execute a tainted EIP for which we could manually correlate to a single CVE that was the same as the one reported by a majority of the AVs in VirusTotal.[3] Thus, for a large majority of the cases, taint-assisted analysis agrees with the AV results. Of the remaining 25 cases, 17 could be identified as misclassifications because the CVE reported by most of the AVs in VirusTotal was not the one that affected the program. The 8 remaining documents were tagged by taint analysis as being false positives even though a CVE was obtained from VirusTotal.

[3] Note that different AVs often tag the same vulnerability with different tags in VirusTotal. We took the tag given by a majority of the reported tags, as the representative of the sample.

C Command and Control (C2) servers

C2	# Emails	C2	# Emails	C2	# Emails	C2	# Emails
61.178.77.169	74	mzyzy.vicp.net	3	www.info-microsoft.com	2	googlehk.dynamicdns.co.uk	1
dtl.dnsd.me	66	mygoodbug.dnsd.info	3	www.uyhanur.nna.cc	2	113.10.201.254	1
ns.dns3-domain.com	55	www.uyghuri.mrface.com	3	www.micosofts.com	2	152.101.38.177	1
dtl.eatuo.com	44	6.test.3322.org.cn	3	100.4.43.2	2	blog.sina.com.cn	1
202.85.136.181	32	218.82.206.229	3	61.234.4.214	1	uyghur.epac.to	1
update.googmail.org	31	uyghur.sov.tw	3	a.yahoohello.com	1	xinxin20080628.gicp.net	1
dtl6.mooo.com	29	3.test.3322.org	3	bc1516.7766.org	1	yah00mail.gicp.net	1
www.discoverypeace.org	26	newwhitehouse.org	3	202.68.226.250	1	hbnjx.6600.org	1
58.64.172.177	22	goodnewspaper.f3322.org	3	msdn.homelinux.org	1	humanbeing2009.gicp.net	1
email.googmail.org	22	nskupdate.com	3	207.204.245.192	1	webhelp01.freetcp.com	1
news.googmail.org	22	webmonder.gicp.net	3	216.131.66.96	1	mobile.yourtrap.com	1
61.128.122.147	17	61.132.74.68	3	www.avasters.com	1	125.141.149.23	1
softmy.jkub.com	15	61.178.77.108	3	202.130.112.231	1	222.73.27.223	1
61.234.4.213	13	betterpeony.com	3	nbsstt.3322.org	1	www.jiapin.org	1
dnsmm.bpa.nu	11	4.test.3322.org	3	goodnewspaper.3322.org	1	ibmcorp.slyip.com	1
121.170.178.221	10	61.234.4.210	3	webposter.gicp.net	1	182.16.11.187	1
zeropan007.3322.org	10	9.test.3322.org.cn	3	uyghur1.webhop.net	1	star2.ksksz.com	1
wwzzsh.3322.org	9	8.test.3322.org.cn	3	webwx.3322.orgxiexie.8866.org	1	69.197.132.130	1
222.77.70.237	9	1.test.3322.org	3	125.141.149.49	1	www.yahooprotect.com	1
3.test.3322.org.cn	8	radio.googmail.org	3	guanshan.3322.org	1	xiexie.8866.org	1
1.test.3322.org.cn	8	7.test.3322.org.cn	3	leelee.dnset.com	1	img.mic-road.com	1
2.test.3322.org.cn	8	tokyo.collegememory.com	2	uygur.eicp.net	1	photo.googmail.org	1
eemete.freetcp.com	8	201.22.184.42	2	kxwss.8800.org	1	tonylee38.gicp.net	1
apple12.crabdance.com	8	61.178.77.96	2	173.208.157.186	1	suggest.dns1.us	1
wolf001.us109.eoidc.net	7	webproxy.serveuser.com	2	rc.arkinixik.com	1	worldview.instanthq.com	1
4.test.3322.org.cn	7	www.bbcnewes.net	2	www.uusuanru.nna.cc	1	goodnewspaper.gicp.net	1
etdt.cable.nu	6	done.youtubesitegroup.com	2	uxz.fo.mooo.com	1	112.121.182.150	1
205.209.159.162	6	alma.apple.cloudns.org	2	uygur.51vip.biz	1	abc69696969.vicp.net	1
br.stat-dns.com	6	webmailsvr.com	2	peopleunion.gicp.net	1	put.adultdns.net	1
66.79.188.23	6	polat.googmail.org	2	free1000.gnway.net	1	loadbook.strangled.net	1
www.southstock.net	6	religion.xicp.net	2	uxz.fo.dnsd.info	1	internet.3-a.net	1
ns1.3322.net	5	connectsexy.dns-dns.com	2	wodebeizi119.jkub.com	1	news.scvhosts.com	1
121.254.173.57	5	dns3.westcowboy.com	2	itsec.eicp.net	1	98.126.20.221	1
www.uyghur.25u.com	5	61.220.138.100	2	stormgo.oicp.net	1	mydeyuming.cable.nu	1
202.96.128.166	5	27.254.41.7	2	boy303.2288.org	1	gshjl.3322.org	1
ns1.oray.net	5	116.92.6.197	2	webjz.9966.org	1	forever001.dtdns.net	1
jhska.cable.nu	5	apple12.co.cc	2	zbing.strangled.net	1	grt1.25u.com	1
test195.3322.org	5	58.64.129.149	2	tommark5454.xxxy.info	1	66.197.202.242	1
61.234.4.218	5	worldmaprsh.com	2	oyghur1.webhop.net	1	kaba.wikaba.com	1
61.128.110.37	5	phinex127.gicp.net	2	addi.apple.cloudns.org	1	221.239.96.180	1
ns1.china.com	5	wxjz.6600.org	2	60.170.255.85	1	174.139.133.58	1
a2010226.gicp.net	5	gecko.jkub.com	2	toolsbar.dns0755.net	1	125.141.149.46	1
logonin.uyghuri.com	4	smtp.126.com	2	61.132.74.113	1	frank.3feet.com	1
macaonews.8800.org	4	errorslog.com	2	113.10.201.250	1	115.126.3.214	1
book.websurprisemail.com	4	uyghurie.51vip.biz	2	home.graffiti.net	1	liveservices.dyndns.tv	1
desk.websurprisemail.com	4	tanmii.gicp.net	2	statistics.netrobots.org	1	inc.3feet.com	1
test.3322.org.cn	4	211.115.207.7	2	freesky365.gnway.net	1	1nsmm.bpa.nu	1
221.239.82.21	4	59.188.5.19	2	greta.ikwb.com	1	www.yahooprotect.net	1
liveservices.dyndns.info	4	206.196.106.85	2	englishclub.2288.org	1	222.82.220.118	1
180.169.28.58	4	religion.8866.org	2	mm.utf888.com	1	webwxjz.3322.org	1
portright.org	4	68.89.135.192	2	annchan.mrface.com	1	61.234.4.220	1
video.googmail.org	4	blogging.blogsite.org	2	www.shine.4pu.com	1	thankyou09.gicp.net	1
www.guzhijiaozihaha.net	4	softjohn.ddns.us	2	copy.apple.cloudns.org	1	218.28.72.138	1
207.46.11.22	4	report.dns-dns.com	2	220.171.107.138	1	soft.epac.to	1
www.googmail.org	4	115.160.188.245	2	uyghuri.mrface.com	1	www.yahooip.net	1
2.test.3322.org	3	newyorkonlin.com	2	218.108.42.59	1	msejake.7766.org	1
dcp.googmail.org	3	tw252.gicp.net	2	58.64.193.228	1	202.67.215.143	1
test.3322.org	3	61.222.31.54	2	tt9c.2288.org	1	www.yahoohello.com	1
np6.dnsrd.com	3	tomsonmartin.ikwb.com	2	forum.universityexp.com	1	202.109.121.138	1