Targeted Threat Index: Characterizing and Quantifying Politically-Motivated Targeted Malware

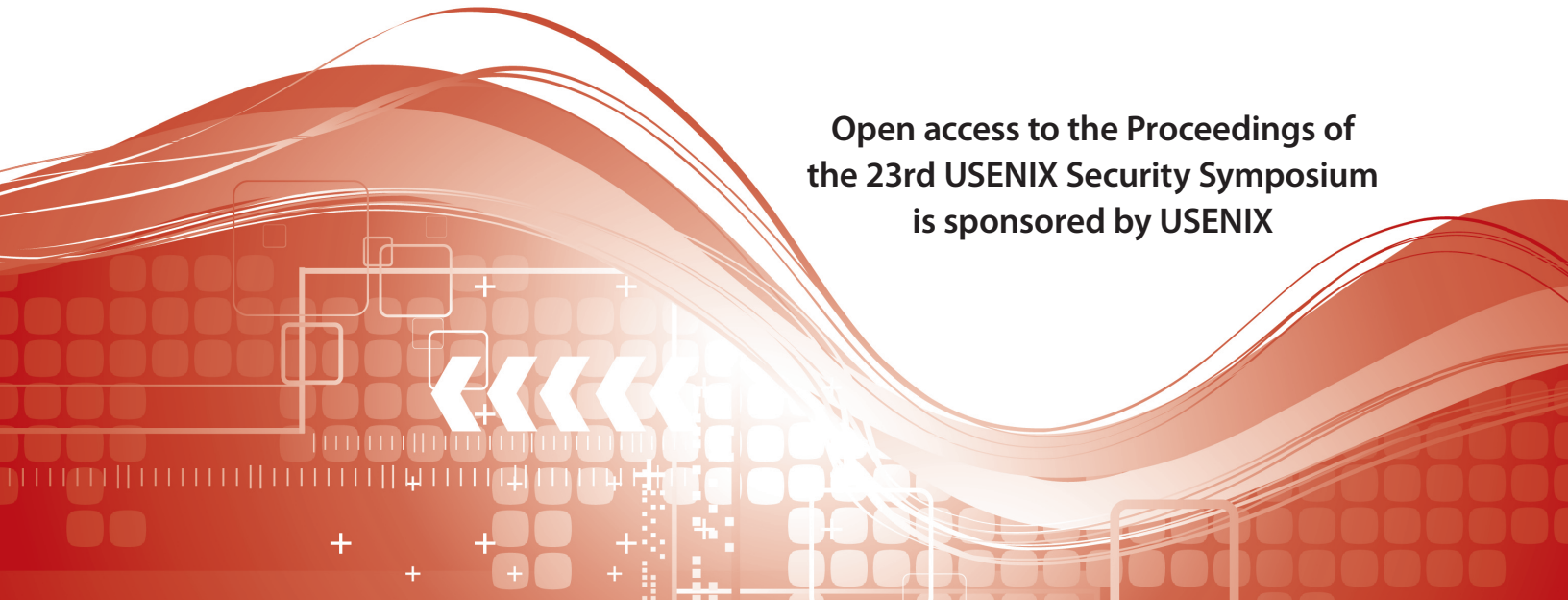
Seth Hardy, Masashi Crete-Nishihata, Katharine Kleemola, Adam Senft, Byron Sonne, and Greg Wiseman, *The Citizen Lab;* Phillipa Gill, *Stony Brook University;* Ronald J. Deibert, *The Citizen Lab*

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

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

Targeted Threat Index: Characterizing and Quantifying Politically-Motivated Targeted Malware

Seth Hardy[§] Masashi Crete-Nishihata[§] Katharine Kleemola[§] Adam Senft[§]
Byron Sonne[§] Greg Wiseman[§] Phillipa Gill[†] Ronald J. Deibert[§]
§ *The Citizen Lab, Munk School of Global Affairs, University of Toronto, Canada*
† *Stony Brook University, Stony Brook, USA*

Abstract

Targeted attacks on civil society and non-governmental organizations have gone underreported despite the fact that these organizations have been shown to be frequent targets of these attacks. In this paper, we shed light on targeted malware attacks faced by these organizations by studying malicious e-mails received by 10 civil society organizations (the majority of which are from groups related to China and Tibet issues) over a period of 4 years.

Our study highlights important properties of malware threats faced by these organizations with implications on how these organizations defend themselves and how we quantify these threats. We find that the technical sophistication of malware we observe is fairly low, with more effort placed on socially engineering the e-mail content. Based on this observation, we develop the Targeted Threat Index (TTI), a metric which incorporates both social engineering and technical sophistication when assessing the risk of malware threats. We demonstrate that this metric is more effective than simple technical sophistication for identifying malware threats with the highest potential to successfully compromise victims. We also discuss how education efforts focused on changing user behaviour can help prevent compromise. For two of the three Tibetan groups in our study simple steps such as avoiding the use of email attachments could cut document-based malware threats delivered through e-mail that we observed by up to 95%.

1 Introduction

Civil society organizations (CSOs), working on human rights issues around the globe, face a spectrum of politically-motivated information security threats that seek to deny (e.g. Internet filtering, denial-of-service attacks), manipulate (e.g. website defacements) or monitor (e.g. targeted malware) information related to their work. Targeted malware attacks in particular are an increasing problem for CSOs. These attacks are not isolated incidents, but waves of attacks organized in campaigns that persistently attempt to compromise systems and gain access to networks over long periods of time while remaining undetected. These campaigns are custom designed for specific targets and are conducted by highly motivated attackers. The objective of these campaigns is to extract information from compromised systems and monitor user activity and is best understood as a form of espionage. CSOs can be particularly susceptible to these threats due to limited resources and lack of security awareness. Targeted malware is an active research area, particularly in private industry. However, focused studies on targeted attacks against CSOs are relatively limited despite the persistent threats they face and the vulnerability of these groups.

In this study, we work with 10 CSOs for a period of 4 years to characterize and track targeted malware campaigns against these groups. With the exception of two groups that work on human rights in multiple countries, the remaining eight groups focus on China and Tibet-related human rights issues. We focus on targeted malware typically delivered via e-mail that is specifically tailored to these groups as opposed to conventional spam which has been well characterized in numerous previous works [27, 42, 45, 52, 70, 71]. We consider the threats to these groups along two axes: the technical sophistication of the malware as well as sophistication of the social engineering used to deliver the malicious payload. We combine these two metrics to form an overall threat ranking that we call the Targeted Threat Index (TTI). While other scoring systems exist for characterizing the level of severity and danger of a technical vulnerability [7, 17, 41, 50], no common system exists for ranking the sophistication of targeted e-mail attacks. TTI allows us to gain insights into the relative sophistication of social engineering and malware leveraged against CSOs.

A key to the success of our study is a unique methodology, combining qualitative and technical analysis of

e-mails and their attachments with fieldwork (e.g. site visits) and interviews with affected CSOs. This methodology, which we describe in more detail in Section 3, allows us to both accurately rate the level of targeting of e-mail messages by interfacing with CSOs participating in our study (Section 4.2), and understand the relative technical sophistication of different malware families used in the attacks (Section 4.3). By combining the strengths of our qualitative and quantitative analysis, we are able to accurately understand trends in terms of social engineering and technical sophistication of politically-motivated targeted malware threats faced by CSOs.

Our study makes the following observations, which have implications for security strategies that CSOs can employ to protect themselves from targeted malware:

Attachments are the primary vector for email based targeted malware. More than 80% of malware delivered to Tibet-related organizations in our study and submitted to us is contained in an e-mail attachment. Further, for 2 of the 3 Tibetan organizations in our study (with at least 40 submitted e-mails), simply not opening attachments would mitigate more than 95% of targeted malware threats that use email as a vector.

Targeted malware technical sophistication is low. Social engineering sophistication is high We find that the technical sophistication of targeted malware delivered to CSOs in our study is relatively low (*e.g.,* relative to commercial malware that has been found targeting CSOs and journalists [35,36,38] and conventional financially motivated malware), with much more effort given to socially engineering messages to mislead users. This finding highlights the potential for education efforts focused on changing user behaviours rather than high-cost technical security solutions to help protect CSOs.

CSOs face persistent and highly motivated actors. For numerous malware samples in our study we observe several versions of the software appearing over the course of our four year study. These multiple versions show evidence of technical improvements to complement existing social engineering techniques.

Since the start of our study we have participated in a series of workshops with the participating Tibetan organizations to translate these results into a training curriculum. Specifically, we have educated them about how to identify suspicious e-mail headers to identify spoofed senders and demonstrated tools that can be used to check e-mailed links for malware and drive-by-downloads.

The rest of the paper is structured as follows. Section 2 presents relevant background on targeted malware and attacks on CSOs. Our data collection methodology is described in Section 3. We describe our targeting and technical sophistication metrics as well as how we combine them to produce the Targeted Threat Index (TTI) in Section 4. Training and outreach implications of our work are discussed in Section 5. We present related work in Section 6 and conclude in Section 7.

2 Background

2.1 Targeted Malware Overview

Targeted malware are a category of attacks that are distinct from common spam, phishing, and financially motivated malware. Spam and mass phishing attacks are indiscriminate in the selection of targets and are directed to the largest number of users possible. Similarly, financially motivated malware such as banking trojans seek to compromise as many users as possible to maximize the potential profits that can be made. The social engineering tactics and themes used by these kinds of attacks are generic and the attack vectors are sent in high volumes. By contrast targeted malware attacks are designed for specific targets, sent in lower volumes, and are motivated by the objective of stealing specific sensitive data from a target.

Targeted malware attacks typically involve the following stages [24, 66]:

Reconnaissance: During this stage attackers conduct research on targets including profiling systems, software, and information security defenses used to identify possible vulnerabilities and contextual information on personnel and activities to aid social engineering.

Delivery: During this stage a vector for delivering the attack is selected. Common vectors include e-mails with malicious documents or links, or contacting targets through instant messaging services and using social engineering to send malware to them. Typically, a target of such an attack receives an e-mail, possibly appearing to be from someone they know, containing text that urges the user to open an attached document (or visit a website).

Compromise: During this stage malicious code is executed on a target machine typically after a user initiated action such as opening a malicious document or link.

Command and Control: During this stage the infected host system establishes a communications channel to a command and control (C&C) server operated by the attackers. Once this channel has been established the attackers can issue commands and download further malware on to the system

Additional attacker actions: After a successful compromise is established, attackers can conduct a number of actions including ex-filtrating data from the infected host and transmitting it back to attackers through a process of encrypting, compressing, and transferring to a server

operated by the attackers. Attackers may also use peripherals such as webcams and microphones to monitor users in real time. The infected host may also serve as a starting point to infect other machines on the network and seek out specific information or credentials.

2.2 Targeted Malware and CSOs

Targeted malware has become recognized by governments and businesses around the world as a serious political and corporate espionage threat. The United States government has been particularly vocal on the threat targeted malware enabled espionage poses. General Keith Alexander, current Director of the National Security Agency and Commander of United States Cyber Command has stated that the theft of US intellectual property through cyber espionage constitutes the "greatest transfer of wealth in history" [47]. Recent widely publicized targeted malware intrusions against Google, RSA, the New York Times and other high profile targets have raised public awareness around these attacks [20, 44, 48]

Despite this increased attention, targeted malware is not a new problem, with over a decade of public reports on these kinds of attacks [66]. However, the majority of research on targeted malware is conducted by private security companies who typically focus on campaigns against industry and government entities. As a result, targeted attacks on civil society and non-governmental organizations have gone underreported despite the fact that these organizations have been shown to be frequently targeted by cyber espionage campaigns. In particular, communities related to ethnic minority groups in China including Tibetans, Uyghurs, and religious groups such as Falun Gong have been frequent targets of cyber espionage campaigns with reports dating back to at least 2002 [61].

In some cases, the same actors have been revealed to be targeting civil society groups, government and industry entities. A notable example of this was the 2009 report by the Citizen Lab, a research group at the University of Toronto, which uncovered the "GhostNet" cyber espionage network. GhostNet successfully compromised prominent organizations in the Tibetan community in addition to 1,295 hosts in 103 countries, including ministries of foreign affairs, embassies, international organizations, and news media [25]. The GhostNet case is not an isolated example, as other reports have shown CSOs (commonly Tibetan organizations) included as targets in campaigns that are also directed to a range of government and industry entities [8, 26, 28, 29, 54–56] Some of these reports include technical details on the CSO specific attacks [26, 28, 54, 55] while others note them as a target but do not address in detail [8, 29, 56].

While the majority of documented targeted malware campaigns against CSOs involve China and Tibet-related groups and potentially China-related attack operators [9–11, 23, 25, 26, 32, 61–65, 67, 68] , these kinds of attacks go beyond China. Recent research and news media have reported attacks against large human rights groups focused on multiple issues and countries [31, 46], and communities related to Syria [18] and Iran [37]. Researchers have also uncovered the use of commercial network intrusion products used to target activists from Bahrain [38], the United Arab Emirates [36], and journalists from Ethiopia [35].

3 Data collection

Since our study involves dealing with e-mail messages which may contain personally identifiable information (PII) and collection of information from CSOs who need to maintain privacy of their data, we consulted with our institutional research ethics board during the design of our study. The methods described below have been submitted to and approved by this board.

3.1 Study Participants

We recruited participants via three main channels: (1) an open call on our Web site, (2) outreach to organizations we had prior relationship with and (3) referrals from participating groups. As part of the study these groups agreed to share technical data (e.g., e-mails with suspicious attachments) and participate in interviews at the onset and end of the study. Their identity and any PII shared with us were kept strictly confidential.

For the purposes of our study, we focused on organizations with missions concerning the promotion or protection of human rights. For purposes of this study, "human rights" means any or all of the rights enumerated under the *Universal Declaration of Human Rights* [60], *the International Covenant on Civil and Political Rights* [58], and the *International Covenant on Economic, Social and Cultural Rights* [59]. We also considered organizations on a case by case basis that have a mission that does not directly implicate human rights, but who may nonetheless be targeted by politically motivated digital attacks because of work related to human rights issues (e.g., media organizations that report on human rights violations).

In total, 10 organizations participated in the study (summarized in Table 1). The majority of these groups work on China-related rights issues and five of these organizations focus specifically on Tibetan rights. The high rate of participation from China and Tibet-related human rights issues is due in part to our previous relationships with these communities and a significant interest and enthusiasm expressed by the groups. In addition to the China and Tibet-related groups, our study also includes

two groups, Rights Group 1 and 2 that work on multiple human rights related issues in various countries.

The majority of organizations operate from small offices with less than 20 employees. Some organizations (China Group 2, Tibet Group 2) have no physical office and consist of small virtual teams collaborating remotely, often from home offices. Of these groups only two (China Group 1, China Group 3) have a dedicated system administrator on staff. Other groups (Tibet Groups 1-5; China Group 2) rely on volunteers or staff with related technical skills (e.g. Web development) to provide technical support. Rights Group 1 and Rights Group 2 are much larger organizations relative to the others in our sample. Both organizations have over 100 employees, multiple offices, dedicated IT teams, and enterprise level computing infrastructures.

3.2 Data Sources

We collect the following pieces of information from the participant groups in order to understand the malware threats they face:

User-submitted e-mail messages. Our primary data source is a collection of e-mails identified by participants as suspicious which were forwarded to a dedicated e-mail server administered by our research team. When available these submissions included full headers, file attachments and / or links. There are three key limitations to relying on user-submitted e-mails for our analysis. First, we are only able to study e-mails identified by participants as suspicious, which may bias our results to only reporting threats that have been flagged by users. Further, individuals may forget to forward e-mails in some cases. Relying on self-reporting also creates bias between groups as individuals at different organizations may have different thresholds for reporting, which creates difficulties in accurately comparing submission rates between groups. Thus the amount of threat behaviour we see should be considered a lower bound on what occurs in practice. Second, having participants forward us e-mails does not allow us to verify if the targeted organization was successfully compromised by the attack (*e.g.,* if another member of the organization open and executed malware on their machine) and what the scope of the attack was. Finally, e-mail is only one vector that may be used to target organizations. Other vectors include waterhole attacks [21], denial of service attacks, or any other vectors (*e.g.,* physical threats like infected USB sticks). These limitations mean that it is possible that we did not comprehensively observe all attacks experienced by our study groups and some more advanced attacks may have gone unreported.

Recognizing the limitations of e-mail submissions, we complement user submitted emails with data from Network Intrusion Detection System (NIDS) alerts, website monitoring, and interviews. Also, upon request of study groups who were concerned of possible infection we analyzed packet capture data from suspect machines. Through the course of this supplementary analysis we did not find indications of malware compromise that used samples that were not included in our pool of user-submitted emails. In this paper we focus on reporting results from analyzing the user submitted emails through the TTI. The NIDS and website monitoring components were added later in our study and do not significantly contribute to TTI analysis. [1]

Table 2: Breakdown of e-mails submitted per group.

Organization Code	# of e-mails
China Group 1	53
China Group 2	18
China Group 3	58
Rights Group 1	28
Rights Group 2	2
Tibet Group 1	365
Tibet Group 2	177
Tibet Group 3	2
Tibet Group 4	97
Tibet Group 5	4

3.3 Overview of User-Submitted E-mails

The e-mails examined in this study span over four years, from October 14, 2009 to December 31, 2013. Data collection began on November 28, 2011, but China Group 3 and Tibet Group 1 forwarded us their pre-existing archives of suspicious emails, resulting in e-mail samples dating back to October 14, 2009. In total, we received 817 e-mails from the 10 groups participating in our study. Table 2 breaks down the submissions from each groups and illustrates that submissions were highly non-uniform across the groups. Thus, in general, we focus on the groups with at least 50 e-mail submissions for our analysis.

Figure 1 shows the cumulative number of e-mail submissions per month over the course of the study. For example, China Group 3 shared a set of e-mails received in 2010 by a highly targeted member of the organization, which can be observed in Figure 1. Tibet Group 1 accounts for the highest number of submissions relative to the other groups due to being one of the first groups in the study and being persistently targeted by politically motivated malware. Tibetan Groups 2 and 4, who joined the study later (in April 2012) show a similar submission rate to original Tibetan Group 1, suggesting these groups are targeted at a similar rate. In Section 4.2, we investi-

Table 1: Summary of groups participating in our study.

Organization Code	Description	Organization size
China Group 1	Human rights organization focused on rights and social justice issues related to China	Small (1-20 employees)
China Group 2	Independent news organization reporting on China	Small (1-20 employees)
China Group 3	Human rights organization focused on rights and social justice issues related to China	Small (1-20 employees)
Rights Group 1	Human rights organization focused on multiple issues and countries	Large (over 100 employees)
Rights Group 2	Human rights organization focused on multiple issues and countries	Large (over 100 employees)
Tibet Group 1	Human rights organization focused on Tibet	Small (1-20 employees)
Tibet Group 2	Human rights organization focused on Tibet	Small (1-20 employees)
Tibet Group 3	Independent news organization reporting on Tibet	Small (1-20 employees)
Tibet Group 4	Human rights organization focused on Tibet	Small (1-20 employees)
Tibet Group 5	Human rights organization focused on Tibet	Small (1-20 employees)

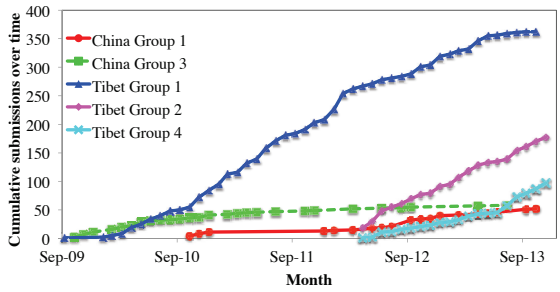


Figure 1: Cumulative number of messages per group over the course of our study for groups that submitted at least 50 e-mail messages.

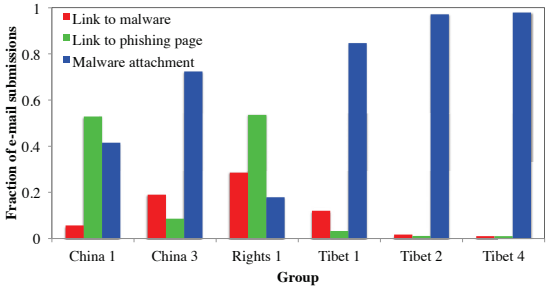


Figure 2: Breakdown of malicious e-mails based on whether they deliver malware as an attachment, refer the use to a link with a malicious file, or attempt to phish data from the user.

gate commonalities in targeting of these groups.

We further classify e-mails as malicious if they include attached malware, a direct link to malware or a site with a drive-by download, or a link to a phishing page. Figure 2 shows the amount of e-mails of each type for the groups that submitted at least 25 e-mails to our system. The most common approach employed in these e-mails was attaching a malicious payload to the e-mail itself. However, we notice a higher rate of phishing attacks on the China-related groups and the rights groups working on multiple international human rights issues. In particular, 46% of the e-mails submitted by China Group 1, and 50% of the e-mails submitted by Rights Group 1, direct the user to a phishing Web site. In the case of China Group 1, this large proportion of phishing sites is observed because this group configured their spam filter to forward e-mails to our system, resulting in us receiving a large number of generic, non-targeted spam. In contrast, the phishing observed for Rights Group 1, while low in volume (13 out of 26 messages) is targeted. We delve more into how we rate the targeting of e-mails in Section 4.2.

The rate of submissions to our project meant that it was feasible to manually analyze e-mail attachments for malware as they were submitted. This analysis gives us higher confidence in our results because AV signatures are frequently unable to detect new or modified threats, and can overlook the presence of a malicious payload that can be easily identified upon manual inspection (e.g. shellcode in an RTF exploit). In total, we analyzed 3,617 payload files and found 2,814 (78%) of them to be malicious. Section 4.3 describes our analysis methodology in more detail.

4 Targeted Threat Index

Our dataset includes a wide range of targeted malware threats varying in level of both social engineering and technical complexity. This range presents a challenge in ranking the relative sophistication of the malware and targeting tactics used by attackers.

While scoring systems such as the Common Vulnerability Scoring System [17] exist for the purpose of communicating the level of severity and danger of a vulnerability, there is no standardized system for ranking

the sophistication of targeted email attacks. This gap is likely because evaluating the sophistication of the targeting is non-technical, and cannot be automated due to the requirement of a strong familiarity with the underlying subject material.

To address this gap we developed the Targeted Threat Index (TTI) to assign a ranking score to the targeted malicious emails in our dataset. The TTI score is intended for use in prioritizing the analysis of incoming threats, as well as for getting an overall idea of how severely an organization is threatened.

The TTI score is calculated by taking a base value determined by the sophistication of the targeting method, which is then multiplied by a value for the technical sophistication of the malware. The base score can be used independently to compare emails, and the combined score gives an indication of the level of effort an attacker has put into individual threats.

4.1 TTI Metric

The TTI score is calculated in two parts:

(*Social Engineering Sophistication Base Value*)
$\times$(*Technical Sophistication Multiplier*) = *TTI Score*

TTI scores range from 1 to 10, where 10 is the most sophisticated attack. Scores of 0 are reserved for threats that are not targeted, even if they are malicious. For example, spam using an attached PDF or XLS to bypass anti-spam filters, and highly sophisticated financially motivated malware, would both score 0.

This section overviews how we compute the *Social Engineering Sophistication Base Value* (Section 4.2) and the *Technical Sophistication Multiplier* (Section 4.3). In Section 4.4, we present the results of computing and analyzing the TTI value of threats observed by the organizations in our study. We also discuss implications and limitations of the metric.

4.2 Social Engineering Tactics

We leverage a manual coding approach to measure the sophistication of social engineering tactics used in the attacks observed by the organizations in our study. While automated approaches may be explored in the future, this manual analysis allows us to have high confidence in our results, especially since understanding the social engineering often required contextual information provided by the organizations in our study. To quantify the level of sophistication, we manually analyse the e-mail subject line, body, attachments and header fields. We perform an initial content analysis by coding the e-mails based on their semantic content, and then use these results to generate a numerical metric quantifying the level of targeting used.

4.2.1 Content coding and analysis results

We code the e-mails based on their subject line, body, attachments and headers using the following methodology:

Subject line, body, and attachments. The content of the subject line, body and attachments for each submitted e-mail were content coded into 8 themes, each containing categories for specific instances of the theme: Country / Region (referring to a specific geographical country or region); Ethnic Groups (referring to a specific ethnic group); Event (referring to a specific event); Organizations (referring to specific organizations); People (referring to specific persons), Political (reference to specific political issues), Technology (reference to technical support), Miscellaneous (content without clear context or categories that do not fall into one of the other themes). Table 3 summarizes the themes and provides examples of categories within each theme.

E-mail headers. The header of each e-mail was analyzed to determine if the sending e-mail address was spoofed or the e-mail address was otherwise designed to appear to come from a real person and / or organization (e.g. by registering an e-mail account that resembles a person and / or organization's name from a free mail provider). We divide the results based on whether they attempted to spoof an organization or a specific person.

Using this manual analysis, we perform a content analysis of e-mails submitted by the organizations. Results of this analysis confirm that social engineering is an important tool in the arsenal of adversaries who aim to deliver targeted malware. Specifically, 95% and 97% of e-mails to Chinese and Tibetan groups, respectively, included reference to relevant regional issues. Spoofing of specific senders and organizations was also prevalent with 52% of e-mails to Tibetan groups designed to appear to come from real organizations, often from within the Tibetan community. For example, a common target of spoofing was the Central Tibetan Administration (CTA), referenced in 21% of the spoofed e-mails, which administers programs for Tibetan refugees living in India and advocates for human rights in Tibet. While the number of e-mail submissions were lower for the general human rights groups, we observe similar trends there with 92% of e-mails submitted by Rights Group 1 appearing to come from individuals in the group (as a result of spoofing).

In some cases we even observed the same attackers targeting multiple CSOs with customized e-mail lures. For example, we tracked a campaign that targeted China Groups 1 and 2, and Tibet Group 1 with a remote access

Table 3: Overview of themes and categories within the themes for grouping targeted e-mail messages.

Theme	Total Categories	Example Categories
Country/Region	26	China, US, European Union
Ethnic Groups	2	Tibetan, Uyghur
Event	31	self immolation, Communist Party of China, 18th National Party Congress
Organizations	32	United Nations, Central Tibetan Administration
People	31	His Holiness the Dalai Lama, Hu Jintao
Political	6	human rights, terrorism
Technology	5	software updates, virtual private servers
Miscellaneous	1	content without clear context which falls outside of the other themes

trojan we call IEXPL0RE [22] China Group 1 received the malware in e-mails claiming to be from personal friends whereas China Group 2 received the malware in an e-mail containing a story about a high-rise apartment building fire in China. In contrast, Tibet Group 1 received the malware embedded into a video of a speech by the Dalai Lama, attached to an e-mail about a year in review of Tibetan human rights issues.

4.2.2 Social Engineering Sophistication Base Value

While the content analysis results clearly show attacks tailored to the interests of targeted groups, content coding alone does not give a relative score of the sophistication used in the attacks. We now describe how we assign the "social engineering sophistication base value" to e-mails based on their level of social engineering.

To measure the targeting sophistication we assign a score that ranges from 0-5 that rates the social engineering techniques used to get the victim to open the attachment. This score considers the content and presentation of the e-mail message as well as the claimed sender identity. This determination also includes the content of any associated files, as malware is often implanted into legitimate relevant documents to evade suspicion from users when the malicious documents are opened.

The Social Engineering Sophistication Base Value is assigned based on the following criteria:

0 Not Targeted: Recipient does not appear to be a specific target. Content is not relevant to the recipient. The e-mail is likely spam or a non-targeted phishing attempt.

1 Targeted Not Customized: Recipient is a specific target. Content is not relevant to the recipient or contains information that is obviously false with little to no validation required by the recipient. The e-mail header and/or signature do not reference a real person or organization.

2 Targeted Poorly Customized: Recipient is a specific target. Content is generally relevant to the target but has attributes that make it appear questionable (e.g. incomplete text, poor spelling and grammar, incorrect addressing). The e-mail header and / or signature may reference a real person or organization.

3 Targeted Customized: Recipient is a specific target. Content is relevant to the target and may repurpose legitimate information (such as a news article, press release, conference or event website) and can be externally verified (e.g. message references information that can be found on a website). Or, the e-mail text appears to repurpose legitimate e-mail messages that may have been collected from public mailing lists or from compromised accounts. The e-mail header and / or signature references a real person or organization.

4 Targeted Personalized: Recipient is a specific target. The e-mail message is personalized for the recipient or target organization (e.g. specifically addressed or referring to individual and / or organization by name). Content is relevant to the target and may repurpose legitimate information that can be externally verified or appears to repurpose legitimate messages. The e-mail header and / or signature references a real person or organization.

5 Targeted Highly Personalized: Recipient is a specific target. The e-mail message is individually personalized and customized for the recipient and references confidential / sensitive information that is directly relevant to the target (e.g. internal meeting minutes, compromised communications from the organization). The e-mail header and / or signature references a real person or organization.

Content coding of emails and determinations of social engineering ratings for the TTI were performed by five independent coders who were given a code book for content categories and the TTI social engineering scale with examples to guide analysis. We performed regular inter-rater reliability checks and flagged any potential edge cases and inconsistencies for discussion and re-evaluation. Following completion of this analysis, two of the authors reviewed the social engineering base value scores to ensure consistency and conformity to the scale. We provide specific examples of each of these targeting values in Appendix A.

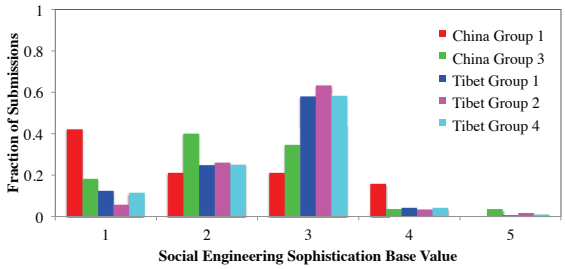


Figure 3: Social engineering sophistication base value assigned to e-mail submissions from groups that submitted at least 50 e-mails.

4.2.3 Summary of Social Engineering Sophistication Base Value

Figure 3 shows the targeting score for organizations in our study who submitted at least 50 e-mails. We can see that actors targeting these groups put significant effort into targeting their messages, in particular the three Tibetan groups included in Figure 3 observe more than half of their messages with a targeting score of 3 or higher. This result means adversaries are taking care to make the e-mail appear to come from a legitimate individual or organization, and include relevant information (*e.g.,* news reports or exchanges from public mailing lists). Higher targeting scores, which result from actions such as personalizing lures to an individual in the group, or including information that requires prior reconnaissance tend to be more rare, but we do observe instances of them. For example, in the case of China Group 3, we observed an e-mail which received a social engineering score of 5, which claimed to be from the group's funder and referenced a specific meeting they had planned that was not public knowledge.

4.3 Technical Sophistication

We manually analyzed all submitted emails and attachments to determine whether they contained politically-motivated malware. The malware is then analyzed in detail to extract information such as the vulnerability, C&C server (if present), and technical sophistication of the exploit.

4.3.1 Assessment methodology

The first step in our analysis pipeline is determining whether the email contains politically motivated malware or not. This process involves an initial inspection for social engineering of the email message and attachment (e.g., an executable pretending to be a document). We also correlate with other emails received as part of this project to identify already-known malware. Well-known malware attacks (*e.g.,* the Zeus trojan masquerading as an email from the ACH credit card payment processor, or Bredolab malware pretending to be from the DHL courier service) are not considered targeted attacks in our study, but are still kept for potential review.

Once we have identified emails which we suspect of containing politically-motivated malware, we perform the following analysis steps on any attachments to verify that they indeed contain malware. First, we run the attachment in a sandboxed VM to look for malicious activity *e.g.,* an Office document writing files to disk or trying to connect to a C&C server. We also check the MD5 hash of the attachment against the Virus Total database to see if it matches existing viruses. We also manually examine the attached file for signs of malicious intent (*e.g.,* executable payload in a PDF, shellcode or Javascript). We exclude any graphics attached to the email which are used for social engineering (and do not contain malicious payload) from our analysis.

We follow this initial analysis with more detailed technical analysis of the attachments which we confirm contain malware. First, we manually verify the file type of the attachment for overview statistics. This manual analysis is necessary as the Unix file command may be misled by methods of manipulating important bytes in the file (*e.g.,* replacing \rtf1 with \rtf[null]). We then identify if the vulnerability included in the malware already exists in a corpus of vulnerabilities, such as the Common Vulnerabilities and Exposures (CVE) naming system. We also perform analysis of network traffic from the attachment to identify the C&C server the malware attempts to contact. In cases where the malware does not execute in our controlled environment we manually examine the file to extract the relevant information.

On a case-by-case basis we use additional tools such as IDA [1] and OllyDbg [3] for detailed static and dynamic analysis, respectively. Our goal in this analysis is to identify relationships between malware campaigns between organizations, or instances of the same malware family repeatedly targeting a given organization. By observing overlapping C&C servers, or mapping malware to common exploits identified by anti virus/security companies we can cluster attacks that we believe come from the same malware family and potentially the same adversary.

4.3.2 Technical Sophistication Multiplier

While the previous analysis is useful for understanding the nature of threats, we also score threats numerically to aid in understanding the relative technical sophistication of their approaches. Each malware sample is assigned one of the following values:

1 Not Protected - The sample contains no code protec-

tion such as packing, obfuscation (e.g. simple rotation of interesting or identifying strings), or anti-reversing tricks.

1.25 Minor Protection - The sample contains a simple method of protection, such as one of the following: code protection using publicly available tools where the reverse method is available, such as UPX packing; simple anti-reversing techniques such as not using import tables, or a call to IsDebuggerPresent(); self-disabling in the presence of AV software.

1.5 Multiple Minor Protection Techniques - The sample contains multiple distinct minor code protection techniques (anti-reversing tricks, packing, VM / reversing tools detection) that require some low-level knowledge. This level includes malware where code that contains the core functionality of the program is decrypted only in memory.

1.75 Advanced Protection - The sample contains minor code protection techniques along with at least one advanced protection method such as rootkit functionality or a custom virtualized packer.

2 Multiple Advanced Protection Techniques - The sample contains multiple distinct advanced protection techniques, e.g. rootkit capability, virtualized packer, multiple anti-reversing techniques, and is clearly designed by a professional software engineering team.

The purpose of the technical sophistication multiplier is to measure how well the payload of the malware can conceal its presence on a compromised machine. We use a multiplier because advanced malware requires significantly more time and effort (or money, in the case of commercial solutions) to customize for a particular target.

We focus on the level of obfuscation used to hide program functionality and avoid detection for the following reasons: (1) It allows the compromised system to remain infected for a longer period; (2) it hinders analysts from dissecting a sample and developing instructions to detect the malware and disinfect a compromised system; (3) since most common used remote access trojans (RATs) have the same core functionality (e.g. keylogging, running commands, exfiltrating data, controlling microphones and webcams, etc.) the level of obfuscation used to conceal what the malware is doing can be used to distinguish one RAT from another.

4.3.3 Summary of Technical Sophistication Multiplier Value

Figure 4 shows the technical sophistication multiplier values for e-mails submitted by the different organizations in our study. One key observation we make here is that the email-based targeted malware that was self-reported by our study groups is relatively simple. The highest multiplier value we see is 1.5 and even that value is seen infrequently. The majority of malware observed is rated either 1 or 1.25 according to our technical scoring criteria, with Tibetan Groups observing a higher fraction of malware rated 1.25 and Chinese groups observing a higher fraction rated 1.

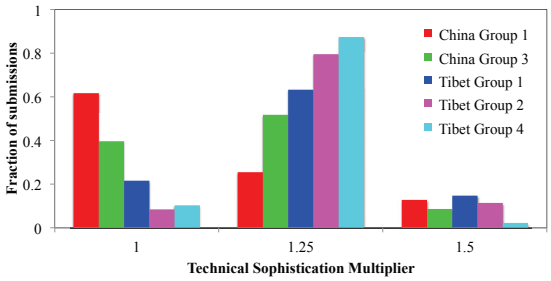


Figure 4: Technical sophistication multiplier assigned to e-mail submissions from groups that submitted at least 50 e-mails.

The technical sophistication multiplier value is also useful for assessing the technical evolution of threats in our study. When we group malware into different family groups we can see some of these groups are under active development. For example, we observe multiple versions of the Enfal [40, 49], Mongal [14], and Gh0st RAT [15] families with increasing levels of sophistication and defenses in place to protect the malware code (resulting in an increase in technical multiplier from 1 to 1.25 for these families). Since our technical multiplier value focuses on how well malware code defends and disguises itself, changes to other aspects of the code may not result in an increase in value (*e.g.,* we observe multiple versions of the IMuler.A/Revir.A malware which all receive a score of 1). Interestingly, when we observe both a Windows and Mac version of a given malware family, the technical score for the Mac version tended to be lower with the Mac version being relatively primitive relative to the Windows variant.

4.4 TTI Results

We now show how the TTI metric can help us better characterize the relative threat posed by targeted malware. Figure 5 shows the technical sophistication multiplier and maximum/minimum TTI scores for malware families observed in our dataset. Since we primarily observe simple malware, with a technical sophistication multiplier of 1 or 1.25, this value does a poor job of differentiating the threat posed by the different malware families to the CSOs. However, by incorporating both the technical sophistication and targeting base value into the TTI metric we can gain more insights into how effective these

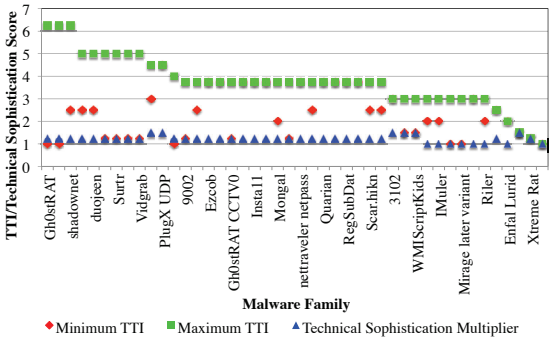


Figure 5: Comparison of the maximum and minimum TTI score and technical sophistication multiplied for malware families observed in our data (sorted in decreasing order of maximum TTI).

Table 4: Top malware families in our data set in terms of technical sophistication multiplier and in terms of final TTI score.

Technical Sophistication		
Family	TTI	Tech. Soph.
3102	3	1.5
nAspyUpdate	1.5	1.5
PlugX	4.5	1.5
PoisonIvy	3	1.5
WMIScriptKids	3	1.5
TTI		
Family	TTI	Tech. Soph. .
Gh0stRAT LURK0	6.25	1.25
shadownet	6.25	1.25
conime	5	1.25
duojeen	5	1.25
iexpl0re	5	1.25

threats may be in practice.

The impact of using TTI is especially apparent when trying to gain insights into the targeted malware that poses the biggest risk to CSOs. Table 4 shows the top 5 malware families we observe in terms of technical sophistication and in terms of TTI score. If we consider the malware families with the highest technical sophistication, we can see that their TTI values are relatively low, with maximums ranging from 1.5 to 4.5. These tend to be malware families that are familiar to researchers. In particular, PlugX and PoisonIvy have been used in targeted attacks together [43] and PlugX is still actively used and under constant development [16]. Despite technical sophistication, the social engineering lures of these threats are not well crafted and pose less of a risk to the CSOs whose members may be able to identify and avoid these threats.

In contrast, the top 5 malware families in terms of TTI have lower technical sophistication (1.25) but much higher levels of social engineering. It is no surprise that threats which score the highest TTI use well known malware that have been extensively documented in attacks against a variety of targets. For example, the TTI scores reflect that Gh0st RAT continues to be seen in higher risk attacks due to its popularity amongst attackers even though it is an older and not particularly advanced tool. Since there is no direct connection between the technical sophistication of threats and the level of social engineering used to target CSOs, it is likely that different threat actors, with a different focus, are at work here. Indeed, Gh0st RAT was discovered by the Citizen Lab in their analysis of GhostNet [25] and IEXPL0RE RAT was discovered and named for the first time in our work.

Another observation is that commercial malware such as FinFisher and DaVinci RCS, while being of much higher technical sophistication (relative to the samples in our study), do not necessarily score higher on TTI than a targeted attack with advanced social engineering and more basic malware. For example, analyzing a FinFisher sample targeted against Bahraini activists [38] with the TTI, produces an overall TTI score that is dependent on the social targeting aspect, even though the malware is very technically advanced. In this case, the FinFisher attack scores 4.0 on the TTI (base targeting score of 2 with a technical multiplier of 2). Although the email used in the attack references the name and organization of a real journalist, the content is poorly customized, and has attributes that look questionable. However, the technical sophistication of the malware is advanced earning it a score of 2 due to multiple advanced protection techniques, including a custom-written virtualized packer, MBR modification, and rootkit functionality. The sample also uses multiple minor forms of protection, including at least half a dozen anti-debugging tricks. Even though the technical multiplier is the maximum value, the overall TTI score is only 4.0 due to the low targeting base value. FinFisher is only effective if it is surreptitiously installed on a users' computer. If the malware is delivered through an email attachment, infection is only successful if the user opens the malicious file. The advanced nature of this malware will cause the overall score to increase quickly with improved targeting, but as it still requires user intervention, this threat scores lower overall than attacks with highly targeted social engineering using less sophisticated malware.

Similar findings can also be observed in attacks using DaVinci RCS developed by Italy-based company Hacking Team against activists and independent media groups from the United Arab Emirates and Morocco [36]. While the malware used in these publicly reported attacks is

technically sophisticated, the social engineering lures employed are poorly customized for the targets resulting in a 4.0 TTI score (targeting base value 2, technical multiplier 2).

These results support the idea that different threat actors have varying focuses and levels of resources, and as a result, different methodologies for attacks. For example, the majority of malware submitted by our study groups appear to be from adversaries that have in-house malware development capabilities and the capacity to organize and implement targeted malware campaigns. These adversaries are spending significant effort on social engineering, but generally do not use technically advanced malware. Conversely, the adversaries using FinFisher and DaVinci RCS have bought these products rather than develop malware themselves. However, while the FinFisher and RCS samples are technically sophisticated pieces of malware, the attacks we analyzed are not sophisticated in terms of social engineering tactics.

4.5 Limitations of TTI

While the Targeted Threat Index gives insight into the distribution of how sophisticated threats are, we are still in the process of evaluating and refining it through interactions with the groups in our study and inclusion of more sophisticated threats observed in related investigations in our lab. Average TTI scores in our dataset may be skewed due to the self-reporting method we use in the study. Very good threats are less likely to be noticed and reported while being sent to far fewer people, and low-quality emails are much more likely to be sent in bulk and stand out. It is also possible that individuals in different groups may be more diligent in submitting samples, which could affect between group comparisons. We are more interested, however, in worst-case (highest) scores and not in comparing the average threat severity between organizations.

Finally, this metric is calculated based on the technical sophistication of the payload, not on the specific exploit. There is currently no method to modify the TTI score in a way similar to the temporal metrics used by the CVSS metric. A temporal metric could be added to increase the final TTI value for 0-day vulnerabilities, or possibly to reduce the score for exploits that are easily detectable due to a public and well-known generation script, e.g. Metasploit [2].

5 Implications

Our study primarily focuses on threats that groups working on human rights issues related to Tibet or China are currently facing. While our dataset is concentrated on these types of groups, our results have implications for how CSOs can protect themselves against email-based targeted malware.

Specifically, we find that moving towards cloud-based platforms (*e.g.,* Google Docs) instead of relying on e-mail attachments would prevent more than 95% of the e-mail malware seen by 2 out of 3 Tibetan groups that had more than 50 e-mail submissions.

Further, our results highlight the potential for lower-cost user education initiatives to guard against sophisticated social engineering attacks, rather than high cost technical solutions. This observation stems from the fact that much of the malware we observe is not technically sophisticated, but rather relies on social engineering to deliver its payload by convincing users to open malicious attachments or links. Other studies [35, 36, 38] that have revealed the use of commercial malware products against CSOs and journalists have shown that many of these cases also rely on duping users into opening malicious e-mail attachments or social engineered instant messaging conversations. These incidents show that even advanced targeted malware requires successful exploitation of users through social engineering tactics.

User education can be a powerful tool against the kinds of targeted attacks we observed in this study. Indeed, the Tibetan community has taken an active approach with campaigns that urge Tibetan users to not send or open attachments and suggests alternative cloud based options such as Google Docs and Dropbox for sharing documents [53]. We have also engaged the Tibetan groups in a series of workshops to introduce training curriculum which draws on examples submitted by organizations participating in our study. We have also provided them with technical background to identify suspicious e-mail headers and how to use free services to check the validity of suspicious links in e-mail messages.

The mitigation strategies presented here are focused on email vectors and do not consider all of the possible attacks these groups may face. We highlight these strategies in particular because the majority of groups in our study identified document-based targeted malware as a high priority information security concern. The adversaries behind these attacks are highly motivated and will likely adapt their tactics as users change their behaviors. For example, it is plausible that if every user in a particular community began to avoid opening attachments and document-based malware infected fewer targets, attackers may move on to vectors such as waterhole attacks or attacks on cloud document platforms to fill the gap. User education and awareness raising activities need to be ongoing efforts that are informed by current research on the state of threats particular communities are experiencing. Evaluation of the effectiveness of user education efforts in at risk communities and corresponding reactions from attackers is required to understand the dynamics between

these processes.

6 Related Work

There is a wide body of literature on filtering and detection methods for spam [27,42,45,52,70,71] and phishing emails and websites [12, 34, 39, 69]. Attention has also been given to evaluating user behavior around phishing attacks and techniques for evading them [6, 30, 33]. By comparison research on detecting email vectors used for targeted malware attacks is limited. A notable exception is [4, 5], which uses threat and recipient features with a random forest classifier to detect targeted malicious emails in a dataset from a large Fortune 500 company. Other work has focused on imporoving detection of documents (e.g. PDF, Microsoft Office) with embedded malicious code [13, 51, 57]

Another area of research explores methods for modeling the stages of targeted attacks and using these models to develop defenses. Guira and Wang [19] propose a conceptual attack model called the attack pyramid to model targeted attacks and identify features that can be detected at the various stages. Hutchins, Cloppert and Amin, [24] use a kill chain model to track targeted attack campaigns and inform defensive strategies.

Metrics have been developed to characterize security vulnerabilities and their severity [7, 41, 50]. The industry standard is the Common Vulnerability Scoring System (CVSS) [17], which uses three metric groups for characterizing vulnerabilities and their impacts. These groups are: base metric group (the intrinsic and fundamental characteristics of a vulnerability that are constant over time and user environments), temporal metric group (characteristics of a vulnerability that change over time but not among user environments) and environmental metric group (characteristics of a vulnerability that are relevant and unique to a particular user's environment). The CVSS is a widely adopted metric, but only rates technical vulnerabilities. Targeted attacks rely on a user action of opening a malicious attachment or visiting a malicious link to successfully compromise a system. Therefore, the sophistication of message lures and other social engineering tactics are an important part of determining the severity of a targeted attack. Systems like the CVSS cannot address this contextual component.

Our study makes the following contributions to the literature. Previous studies of targeted attacks against CSOs usually focus on particular incidents or campaigns and do not include longitudinal observations of attacks against a range of CSO targets. While standards exist for rating the sophistication of technical vulnerabilities and research has been done on detecting targeted malware attacks and modeling campaigns, there is no scoring system that considers both the sophistication of malware and social engineering tactics used in targeted malware attacks. We address this gap through development of the TTI and validate the metric against four years of data collected from 10 CSOs.

7 Conclusions

Our study provides an in-depth look at targeted malware threats faced by CSOs. We find that considering the technical sophistication of these threats alone is insufficient and that educating users about social engineering tactics used by adversaries can be a powerful tool for improving the security of these organizations. Our results point to simple steps groups can take to protect themselves from document-based targeted malware such as shifting to cloud-based document platforms instead of relying on attachments which can contain exploits. Further research is needed to measure the effectiveness of education strategies for changing user behaviour and how effective these efforts are in mitigation of document-based malware for CSOs. Further work is also required in monitoring how attackers adapt tactics in response to observed behavioural changes in targeted communities.

In ongoing work we are continuing our collection of e-mails and NIDS alerts as well as monitoring other attacks against these groups (*e.g.,* waterhole attacks and DoS attacks) to understand how threats vary based on their delivery mechanism. We are also working to extend our methodology to more diverse CSO communities such as those in Latin America, Africa, and other underreported regions to better document the politically motivated digital threats they may be experiencing.

Acknowlegements

This work was supported by the John D. and Catherine T. MacArthur Foundation. We are grateful to Jakub Dalek, Sarah McKune, and Justin Wong for research assistance. We thank the USENIX Security reviewers and our shepherd Prof. J. Alex Halderman for helpful comments and guidance. We are especially grateful to the groups who participated in our study.

References

[1] https://www.hex-rays.com/products/ida/.

[2] http://www.metasploit.com/.

[3] http://www.ollydbg.de/.

[4] AMIN, R. M. *Detecting Targeted Malicious Emails Through Supervised Classification of Persistent Threat and Recipient Oriented Features*. Doctor of philisophy, George Washington University, 2011.

[5] AMIN, R. M., RYAN, J., H, J. C., AND VAN DORP, J. R. Detecting Targeted Malicious Email. *IEEE Security & Privacy 10*, 3 (2012), 64–71.

[6] BLYTHE, M., PETRIE, H., AND CLARK, J. A. F for Fake: Four Studies on How We Fall for Phish. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2011), CHI '11, ACM, pp. 3469–3478.

[7] CERT. Vulnerability Notes Database Field Descriptions, 2014.

[8] CHIEN, E., AND O'GORMAN, G. The Nitro Attack: Stealing Secrets from the Chemical Industry. Tech. rep., Symantec, 2011.

[9] CITIZEN LAB. Information Operations and Tibetan Rights in the Wake of Self-Immolations: Part I. Tech. rep., University of Toronto, 2012.

[10] CITIZEN LAB. Recent Observations in Tibet-Related Information Operations: Advanced social engineering for the distribution of LURK malware. Tech. rep., University of Toronto, 2012.

[11] CITIZEN LAB. Permission to Spy: An Analysis of Android Malware Targeting Tibetans. Tech. rep., University of Toronto, 2013.

[12] COVA, M., KRUEGEL, C., AND VIGNA, G. There is no free phish: an analysis of free and live phishing kits. In *Proceedings of the 2nd Conference on USENIX Workshop on Offensive Technologies* (July 2008), USENIX Association, p. 4.

[13] CROSS, J. S., AND MUNSON, M. A. Deep pdf parsing to extract features for detecting embedded malware. Tech. rep., Sandia National Laboratories, 2011.

[14] DEEP END RESEARCH. Library of Malware Traffic Patterns, 2013.

[15] FAGERLAND, S. The Many Faces of Gh0st Rat. Tech. rep., Norman, 2012.

[16] FAGERLAND, S. PlugX used against Mongolian targets. Tech. rep., 2013.

[17] FIRST. Common Vulnerability Scoring System (CVSS-SIG), 2007.

[18] GALPERIN, EVA, MARQUIS-BOIRE, MORGAN, SCOTT-RAILTON, J. Quantum of Surveillance: Familiar Actors and Possible False Flags in Syrian Malware Campaigns — Electronic Frontier Foundation. Tech. rep., Electronic Frontier Foundation and The Citizen Lab, University of Toronto.

[19] GIURA, P., AND WANG, W. A Context-Based Detection Framework for Advanced Persistent Threats. *International Conference on Cyber Security (CyberSecurity) 0* (2012), 69–74.

[20] GOOGLE. A new approach to China, 2012.

[21] GRAGIDO, W. Lions at the Watering Hole: The VOHO Affair. Tech. rep., RSA, 2012.

[22] HARDY, S. IEXPL0RE RAT. Tech. rep., Citizen Lab, University of Toronto, 2012.

[23] HARDY, SETH KLEEMOLA, K. Surtr: Malware Family Targeting the Tibetan Community. Tech. rep., Citizen Lab, University of Toronto, 2013.

[24] HUTCHINS, E. M., CLOPPERT, M. J., AND AMIN, R. M. Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. In *6th International Conference on Information Warfare and Security* (2011).

[25] INFORMATION WARFARE MONITOR. Tracking GhostNet: Investigating a Cyber Espionage Network. Tech. rep., University of Toronto, 2009.

[26] INFORMATION WARFARE MONITOR. Shadows in the Cloud: Investigating Cyber Espionage 2.0. Tech. rep., University of Toronto, 2010.

[27] JUNG, J., AND SIT, E. An Empirical Study of Spam Traffic and the Use of DNS Black Lists. In *Proceedings of the 4th ACM SIGCOMM Conference on Internet Measurement* (New York, NY, USA, 2004), IMC '04, ACM, pp. 370–375.

[28] KASPERSKY LAB. The NetTraveler Attacks. Tech. rep., Trend Micro, 2013.

[29] KASPERSKY LAB. Unveiling "Careto" - The Masked APT. Tech. rep., 2014.

[30] KIRLAPPOS, I., AND SASSE, M.-A. Security Education against Phishing: A Modest Proposal for a Major Rethink. *Security Privacy, IEEE 10*, 2 (Mar. 2012), 24–32.

[31] KREBS, B. Espionage Hackers Target Watering Hole Sites, 2012.

[32] LI, F., LAI, A., AND DDL, D. Evidence of Advanced Persistent Threat: A case study of malware for political espionage. In *2011 6th International Conference on Malicious and Unwanted Software* (Oct. 2011), IEEE, pp. 102–109.

[33] LIN, E., GREENBERG, S., TROTTER, E., MA, D., AND AYCOCK, J. Does Domain Highlighting Help People Identify Phishing Sites? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2011), CHI '11, ACM, pp. 2075–2084.

[34] MAIORCA, D., CORONA, I., AND GIACINTO, G. Looking at the Bag is Not Enough to Find the Bomb: An Evasion of Structural Methods for Malicious PDF Files Detection. In *Proceedings of the 8th ACM SIGSAC Symposium on Information, Computer and Communications Security* (New York, NY, USA, 2013), ASIA CCS '13, ACM, pp. 119–130.

[35] MARCZAK, B., GUARNIERI, C., MARQUIS-BOIRE, M., AND SCOTT-RAILTON, J. Hacking Team and the Targeting of Ethiopian Journalists. Tech. rep., Citizen Lab, University of Toronto, 2014.

[36] MARQUIS-BOIRE, M. Backdoors are Forever: Hacking Team and the Targeting of Dissent. Tech. rep., Citizen Lab, University of Toronto, 2013.

[37] MARQUIS-BOIRE, M. Iranian anti-censorship software 'Simurgh' circulated with malicious backdoor. Tech. rep., Citizen Lab, University of Toronto, 2013.

[38] MARQUIS-BOIRE, M., MARCZAK, B., GUARNIERI, C., AND SCOTT-RAILTON, J. For Their Eyes Only: The Commercialization of Digital Spying. Tech. rep., Citizen Lab, University of Toronto, 2013.

[39] MAURER, M.-E., AND HÖFER, L. Sophisticated Phishers Make More Spelling Mistakes: Using URL Similarity against Phishing. In *CSS* (2012), pp. 414–426.

[40] MCAFEE. Enfal, 2008.

[41] MICROSOFT CORPORATION. Security Bulletin Severity Rating System, 2012.

[42] PANTEL, P., AND LIN, D. SpamCop: A Spam Classification & Organization Program. In *Learning for Text Categorization: Papers from the 1998 Workshop* (1998), pp. 95–98.

[43] PAZ, R. D. PlugX: New Tool For a Not So New Campaign. Tech. rep., Trend Micro, 2012.

[44] PERLROTH, N. Chinese Hackers Infiltrate New York Times Computers, Jan. 2013.

[45] RAMACHANDRAN, A., FEAMSTER, N., AND VEMPALA, S. Filtering spam with behavioral blacklisting. In *Proceedings of the 14th ACM conference on Computer and communications security - CCS '07* (New York, New York, USA, Oct. 2007), ACM Press, p. 342.

[46] RILEY, M., AND LAWRENCE, D. Hackers Linked to Chinas Army Seen From EU to D.C., 2012.

[47] ROGIN, J. NSA Chief: Cybercrime constitutes the greatest transfer of wealth in history. *Foreign Policy* (2012).

[48] RSA. Anatomy of an Attack.

[49] SANCHO, DAVID, VILLENEUVE, N. LURID: Attribution Isn't Easy. Tech. rep., Trend Micro.

[50] SANS. @Risk: The Consensus Security Alert, 2014.

[51] SMUTZ, C., AND STAVROU, A. Malicious PDF Detection Using Metadata and Structural Features. In *Proceedings of the 28th Annual Computer Security Applications Conference* (New York, NY, USA, 2012), ACSAC '12, ACM, pp. 239–248.

[52] TAYLOR, B. Sender Reputation in a Large Webmail Service. In *Third Conference on Email and Anti-Spam (CEAS 2006)* (2006).

[53] TIBET ACTION INSTITUTE. https://tibetaction.net/detach-from-attachments/.

[54] TREND MICRO. IXESHE: An APT campaign. Tech. rep., 2012.

[55] TREND MICRO. Luckycat Redux: Inside an APT campaign with multiple targets in India and Japan. Tech. rep., 2012.

[56] TREND MICRO. 2Q Report on Targeted Attack Campaigns. Tech. rep., 2013.

[57] TZERMIAS, Z., SYKIOTAKIS, G., POLYCHRONAKIS, M., AND MARKATOS, E. P. Combining Static and Dynamic Analysis for the Detection of Malicious Documents. In *Proceedings of the Fourth European Workshop on System Security* (New York, NY, USA, 2011), EUROSEC '11, ACM, pp. 4:1—-4:6.

[58] UNITED NATIONS. International Covenant on Civil and Political Rights.

[59] UNITED NATIONS. International Covenant on Economic, Social and Cultural Rights.

[60] UNITED NATIONS. The Universal Declaration of Human Rights.

[61] VAN HORENBEECK, M. Crouching PowerPoint, Hidden Trojan. In *24th Chaos Communications Congress* (2007).

[62] VAN HORENBEECK, M. Cyber attacks against Tibetan communities. Tech. rep., Sans Institute, 2008.

[63] VAN HORENBEECK, M. Is Troy Burning? An overview of targeted trojan attacks. In *SANSFire 2008* (2008).

[64] VILLENEUVE, N. Human Rights and Malware Attacks. Tech. rep., Citizen Lab, University of Toronto, 2010.

[65] VILLENEUVE, N. Nobel Peace Prize, Amnesty HK and Malware. Tech. rep., Citizen Lab, University of Toronto, 2010.

[66] VILLENEUVE, N. Trends in targeted attacks. Tech. rep., Trend Micro, 2011.

[67] VILLENEUVE, N., AND WALTON, G. Targeted Malware Attack on Foreign Correspondents based in China. Tech. rep., Information Warfare Monitor, University of Toronto, 2009.

[68] VILLENEUVE, N., AND WALTON, G. 0day: Civil Society and Cyber Security. Tech. rep., Information Warfare Monitor, University of Toronto, 2009.

[69] XIANG, G., HONG, J., ROSE, C. P., AND CRANOR, L. CANTINA+: A Feature-Rich Machine Learning Framework for Detecting Phishing Web Sites. *ACM Trans. Inf. Syst. Secur. 14*, 2 (Sept. 2011), 21:1—-21:28.

[70] ZHANG, L., ZHU, J., AND YAO, T. An Evaluation of Statistical Spam Filtering Techniques. *Transactions on Asian Language Information Processing 3*, 4 (Dec. 2004), 243–269.

[71] ZHOU, Y., MULEKAR, M. S., AND NERELLAPALLI, P. Adaptive Spam Filtering Using Dynamic Feature Space. In *Proceedings of the 17th IEEE International Conference on Tools with Artificial Intelligence* (Nov. 2005), IEEE Computer Society, pp. 302–309.

```
From:  world fdc <fdc2008paris@gmail.com>
To:  [Tibet Group 1]
Subject:  Invitation

Please reply

1 Attachment:  invitation.doc
```

Figure 6: Example of e-mail with Targeting Score 1

```
From:  ciran nima <nimaciran@gmail.com>
To:  [Tibet Group 1]
Date:  18 Aug 2011
Subject:  Truth of monk dies after setting
himself on fire

Truth of monk dies after setting himself on
fire

1 Attachment:  Truth of monk dies after
setting himself on fire.doc
```

Figure 7: Example of e-mail with Targeting Score 2

Notes

[1] We report on results from other collection sources (e.g. NIDS alerts, website monitoring, and interviews), and cluster analysis of campaigns in a forthcoming technical report available at https://citizenlab/targeted-threats

Appendix

A Examples of targeted e-mails

In this section, we provide specific examples of e-mails that would be assigned targeting scores described in Section 4.2.2.

Targeting Score 1 (Targeted Not customized). The e-mail in Figure 6 was sent to Tibet group 1. The message content and sender are vague and do not relate to the interest of the group. The attachment is a word document implanted with malware. The lack of relevant information in this message gives it a score of 1 (targeted, not customized).

Targeting Score 2 (Targeted, Poorly Customized). The e-mail in Figure 7 was sent to Tibet group 1. It references Tibetan self-immolations which is an issue of interest to the group. However, the sender does not appear to be from a real person or organization. The message content is terse and does not referenced information that can be externally validated. Therefore this message scores a 2 (targeted, poorly customized).

```
From:  Palden Sangpo
<palden.sangpo@tibetancareers.org>
Subject:  Activity Report from Tibetan
Career Centre, Bylakuppe
Date:  24 Jan 2013
To:  [Tibet Group 2]

Dear Sir/Madam,

Tashi Delek.

Please find the attachment of the activity
report of Tibetan Career Centre, Bylakuppe
with this mail.  As I was asked to send
this activity report to your office.

Thank you.

Regards,
Palden Sangpo, Consultant.
Tibetan Career Centre,
Old Guest House, Lugsam Tibetan Settlement
Office,
PO Bylakuppe, Mysore District, Karnataka
State - 571 104
E-mail:  palden.sangpo@tibetancareers.org,
MO +91 9901407808, Off +91 8971551644
www.tibet.jobeestan.com

1 Attachment:  Report to CTA home.doc
```

Figure 8: Example of e-mail with Targeting Score 3

```
From:  Cheng Li <chengli.brookings@aol.com>
Subject:Happy Tib Losar and Ask You a
Favour
23 Feb 2012
To:  [Tibet Group 1]

Dear [Redacted]

I am Cheng Li from John L. Thornton China
Center of Brookings.  I will attend an
annual meeting on Religious Research
with CIIS in Shanghai next week, and
plan to take the chance to visit Tibet.
Attached is a list of tibetans who have
self-immolated from 2009 which my assistant
prepared for me, but i am not sure of its
accuracy.  Would you please have a look
and make necessary corrections.  I will be
really much appreciated if you could do me
the favor and offer some more information
about the latest happenings inside tibet.

Thank you again and happy Tib losar!

Cheng Li
Director of Research, John L. Thornton
China Center
Brookings Institution

1 Attachment:  list_of_self_immolations.
xls
```

Figure 9: Example of e-mail with Targeting Score 4

Targeting Score 3 (Targeted Customized). The e-mail in Figure 8 was sent to Tibet group 2. On the surface it appears to be a professional e-mail from "Palden Sangpo" a consultant at the Tibet Career Centre. The e-mail sender address and signature reference accurate contact details that can be easily verified through an Internet search. However, the e-mail headers reveal the purported e-mail sender address is fraudulent and the actual sender was `albano_kuqo@gmx.com`. The e-mail generally addresses the organization rather than the individual recipient. Therefore this message scores a 3 (targeted, customized).

Targeting Score 4 (Targeted Personalized). The e-mail in Figure 9 was sent to Tibet group 1. It is directly addressed to the director of the group and appears to come from Mr. Cheng Li, a prominent China scholar based at the Brookings Institute. The e-mail address is made to appear to be from Mr. Cheng Li, but from an AOL account (chengli.brookings@aol.com) that was registered by the attackers. The message asks the recipient for information on recent Tibetan self-immolations. The level of customization and personalization used in this message gives it a score of 4 (targeted, personalized).

Targeting Score 5 (Targeted Highly Personalized). Targeting scores of 5 (targeted, highly personalized) require reference to internal information to the target organization that could *not be* obtained through open sources. Examples of messages scoring at this level include an e-mail that purported to come from a funder of China Group 3 that provided details of an upcoming meeting the group actually had scheduled with the funder. In another example, Tibet Group 2 and Tibet Group 3 received separate e-mails that contained specific personal details about a South African group's visit to Dharamsala, India that appear to have been repurposed from a real private communication. The malicious attachment contained an authentic travel itinerary, which would be displayed after the user opened the document. The private information used in these messages suggest that the attackers performed significant reconnaissance of these groups and likely obtained the information through prior compromise.