Citizen Lab, Munk School of Global Affairs, University of Toronto

MUNK SCHOOL OF GLOBAL AFFAIRS

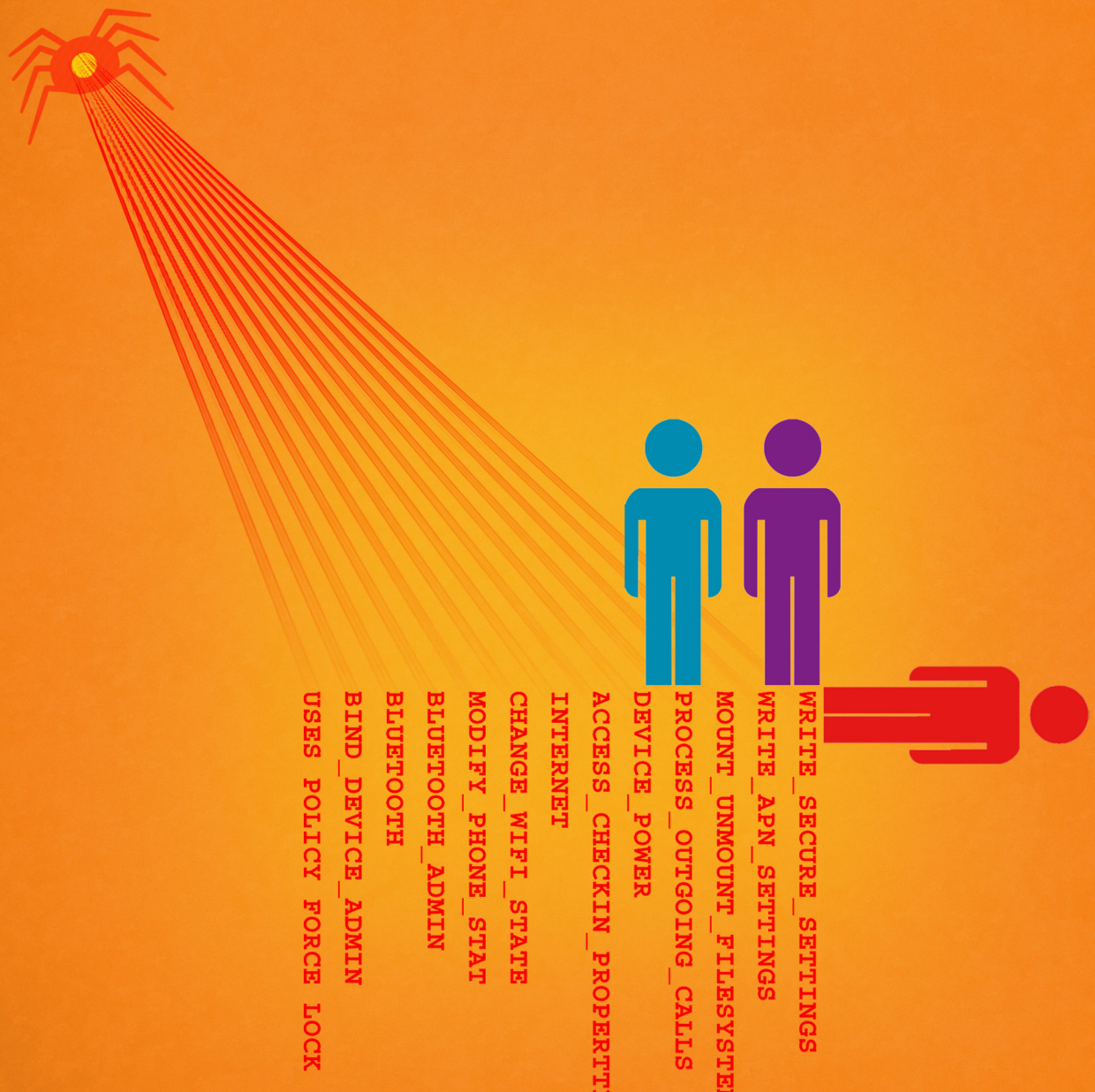
UNIVERSITY OF TORONTO

COMMUNITIES @ RISK

Targeted Digital Threats Against Civil Society

November 11, 2014

https://targetedthreats.net/



EXTENDED ANALYSIS:

2.1

Summary, Methodology, and Data Overview

Summary

Communities @ Risk: Targeted Digital Threats Against Civil Society reports on an intensive study that analyzes targeted digital threats against 10 civil society organizations (CSOs) over a period of four years.

The report combines two major sections:

1. The **Executive Summary** provides detail on how the study was organized and why we feel it is important to read, a high-level overview of the key findings of the research, and considerations about next steps for several stakeholder communities in responding to targeted digital threats.
2. The **Extended Analysis** explains our methodology, and examines the detailed data we gathered during the study period. It is the evidentiary basis for the claims we make in the Executive Summary, and will likely be of interest to a more specialized audience (although we hope everyone will read it).

KEY FINDINGS

In the Executive Summary, we outline five high-level findings. We summarize them again below, while adding more granular details that are given extended treatment in the analysis that follows.

In the digital realm, CSOs face the same threats as the private sector and government, while equipped with far fewer resources to secure themselves.

Through cluster analysis we identify 10 distinct targeted malware campaigns. We find that five of these campaigns have connections to threat actors, previously reported to have targeted government and private industries. CSOs have limited resources and technical capacity, which makes responding to threats a challenge. We generally find that, due to resource constraints, CSOs focus their digital security strategies on user education and behavioural change rather than expensive technical solutions.

Counterintuitively, technical sophistication of malware used in these attacks is low, but the level of social engineering employed is high.

We develop the Targeted Threat Index, a metric for quantifying and characterizing the sophistication of targeted malware attacks. Using this metric, we find that the technical sophistication of targeted malware delivered to CSOs in our study is relatively low (e.g., relative to commercial "lawful intrusion" surveillance kits and conventional financially motivated malware), with much more effort given to socially engineering messages to mislead users.

Digital attacks against CSOs are persistent, adapting to targets in order to maintain access over time and across platforms.

Our analysis of attacks against CSOs over four years allows us to track how attackers change tactics. For numerous malware samples, we observe several versions of the malware appearing over the course of our study. These multiple versions show evidence of technical improvements to complement increasingly refined social engineering techniques. In some cases, we observe threat actors quickly changing tactics to adapt to shifting platform adoption and user behaviour.

Targeted digital threats undermine CSOs' core communications and missions in a significant way, sometimes as a nuisance or resource drain, more seriously as a major risk to individual safety.

The impact of targeted digital attacks against technical systems is apparent and receives ample attention from researchers. However, we find evidence of wider impacts that are not always as obvious, including psychosocial strain and possible connections to physical harms (e.g., arrest and detention). Tracing connections between compromises and harm is challenging, because the relationship between digital compromises and the use of the compromised information by threat actors is indirect. Unlike the consequences of physical threats, which are often readily observable, the most serious impacts of digital threats are typically at least one step removed from the technology that has been exploited.

Targeted digital threats extend the "reach" of the state (or other threat actors) beyond borders and into "safe havens."

The ways CSOs develop their perceptions of risk and threat stemming from targeted attacks depend in part on the physical proximity of their threat actor. Groups operating within the jurisdiction of a repressive regime have greater concerns over physical security and other direct interference from authorities. Conversely, groups situated

outside of a physical jurisdiction controlled by an adversary may prioritize digital threats over physical threats. For groups in diaspora and exile communities, targeted digital threats can be seen as a means for a powerful threat actor, such as a state, to extend their reach beyond borders and into "safe areas."

EXTENDED ANALYSIS STRUCTURE

The Extended Analysis is structured into the following three sections. Each of these sections can be downloaded individually or read as a whole.

Summary, Methodology, and Data Overview outlines our mixed methods approach which incorporates analysis of technical and contextual data using methodologies from the field of information security and the social sciences, and presents a high level overview of our dataset.

Cluster Analysis provides detailed technical analysis of 10 distinct targeted malware campaigns.

Civil Society Perspectives and Responses reports on results from interview data and is a window into how groups under threat think about and respond to digital threats.

We also are publishing data that provide indicators of compromise (including YARA signatures of malware families, MD5 hashes of samples, and command-and-control servers), which are available on our github account and accessible through our project website.

Methodology

This section describes our methodology for data collection and analysis. Since our study involves the collection of potentially sensitive information from civil society organizations, and requires us to deal with personally identifiable information (PII), we consulted with the University of Toronto's Research Ethics Review Board during the design of our study. The methods described below have been submitted to and approved by this board.

STUDY PARTICIPANTS

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

Organizations with a mission concerning the promotion or protection of human rights were eligible to participate.[1] We also considered, on a case-by-case basis, organizations with a mission that does not directly address human rights, but which may engage in work related to human rights issues (e.g., media outlets that regularly report on human rights violations).

In total, 10 organizations participated in the study. The majority of these groups work on China-related rights issues, and five of these organizations focus specifically on Tibetan rights. The exceptions to the China- / Tibet-focused groups in our study are two large organizations that work on multiple human rights-related issues in various countries.

1 For purposes of this study, "human rights" means any or all of the rights enumerated under the Universal Declaration of Human Rights; the International Covenant on Civil and Political Rights; and the International Covenant on Economic, Social and Cultural Rights.

TABLE 1: Study organizations

ORGANIZATION CODE	DESCRIPTION	ORGANIZATION SIZE
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)
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)
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)

Tibet Groups

Dharamsala is a small city in northern India set on the foothills of the Himalayas. His Holiness the Dalai Lama (HHDL) has lived in Dharamsala since 1959 following his escape from Tibet. Dharamsala is the base of the Central Tibetan Administration, which administers programs and schools for Tibetan refugees living in India and advocates for the rights of Tibetans in Tibet. It is also home to many Tibetan NGOs and independent media groups, and thousands of Tibetan refugees. This high concentration of prominent Tibetan institutions makes Dharamsala a prime target for malware campaigns. It has been called one of the most hacked places in the world. For exiled Tibetans, this heightened level of digital risk compounds the many challenges of living as refugees in a developing country.

Three of the Tibet Groups in our study are headquartered in Dharamsala, and two

maintain regional offices there. Across these groups participants expressed challenges related to awareness of threats, low resources, and limited technical capacities.

Tibet Groups reported varying levels of awareness of digital risks in the community. While many participants noted that security awareness was generally increasing among Tibetans, others cautioned that some groups still do not have policies or response plans around targeted digital attacks and "continue to back burn things like security."[2]

A major challenge identified by the Tibet Groups is a lack of technical capacity and resources in the community. Most Tibetan NGOs do not have dedicated system administrators. In some groups, staff members responsible for web development also take on double duty as system administrators. In addition to local staff, there are transient volunteers who come into the community to help with technical projects. As one of these volunteers noted, however, when volunteers leave the community projects sometimes end up unmaintained or completely abandoned.

While the unique circumstances of the Tibetan exile community are challenging, some groups are also taking proactive measures to increase digital security awareness. For example, one of our participating organizations prioritizes digital security in the community within its mission, focusing on raising awareness and user education. These grassroots initiatives demonstrate a growing commitment to addressing security challenges, despite ever-present resource limitations.

China Groups

The three China Groups all work on issues related to human rights and politics in China, but from outside of mainland China. China Groups 1 and 3 each have a central office and one regional branch. China Group 2 operates an independent news website from an office with limited staff. China Group 1 has a program manager that oversees technical projects, but does not have a dedicated system administrator on staff. Instead the group outsources management of its information technology infrastructure to a private company. China Group 3 has had a dedicated system administrator since its founding.

The work of these groups is politically sensitive and has attracted attention from Chinese authorities. China Groups 1 and 2 especially have come under pressure for

2 Tibet Group 1, Program Director, 2011

human rights advocacy and the dissemination of sensitive news, respectively. As China Group 1 explained, "Chinese authorities ... have very clearly in public designated us as an anti-China organization."

These groups are all highly aware of targeted digital threats, and have experienced numerous prior incidents. All of the groups had received targeted malware in the past and their websites are consistently blocked in China. The website of China Group 2 has been repeatedly hit by distributed denial-of-service attacks.

Rights Groups

Rights Groups 1 and 2 are much larger organizations relative to the others in our study. Both have over 100 employees, multiple offices, enterprise level computing infrastructures, and dedicated IT teams and support desks.

These groups act as hub organizations. Rights Group 1, for example, supports multiple regional offices and CSO partners around the world. Rights Group 2, similarly, operates regional branches and is responsible for a large group of staff operating in numerous field locations.

Both groups contend with securing their head offices and maintaining awareness of threats faced by field offices. These challenges show that while the Rights Groups have greater resources they must grapple with a potentially wider spectrum of threats in multiple contexts and countries.

DATA SOURCES

Email Submissions: The majority of data collected consisted of emails identified by participants as suspicious, which were forwarded to a dedicated mail server administered by our research team. When available, these submissions included full headers, file attachments, and / or links.

Relying on forwarded emails presents a collection bias as the recipients must be able to identify that the emails are suspicious and remember to forward the samples to our research team. This collection method also limits the threats studied to those that are sent over email. Additionally, collection of forwarded email samples does not allow us to verify if a targeted organization was successfully compromised by an attack, or the scope of the attack. Recognizing this limitation, we added two more data collection methods to complement the collection of emails.

Network Intrusion Detection System: As an optional study component, we offered to install a network intrusion detection system (NIDS) inside the networks of the participants. In total, seven groups opted into the NIDS project. We used a combination of community and commercial rulesets, as well as a set of custom rules based on threats we analyzed from the email submissions. By placing a NIDS inside an organization's network, we were able to record incoming threats using vectors other than email, as well as detect and observe systems that had already been compromised.

Website Monitoring: We conducted external scans of the study organizations' websites to monitor for potential compromises such as watering hole attacks. These scans were done with publicly available tools including Cyberspark and URL Query.

Interviews and Fieldwork: To gain insights into the experiences of our groups, we conducted a series of semi-structured interviews over a four-year period and made site visits to their offices and locales. While there have been previous technical studies on targeted threats affecting CSOs, it is rare that the context surrounding these attacks and the experiences of the people facing them are properly explored. Interviews and site visits help provide insight into these vital elements.

When possible we conducted interviews with a senior staff member responsible for organizational programming (e.g., executive director, program manager), and a staff member responsible for technical support (e.g., system administrator, webmaster). The interviews explored the organizations' uses of and policies around technology, perceptions of digital security and threats, responses to threats, and the impact of threats. These interviews, coupled with site visits and participant observations, helped us understand the working conditions, routines, infrastructure, and local social and political context that form the day-to-day environment of our participants.

Interviews were held opportunistically and did not follow a set schedule. The total number of interviews per group is outlined in Table 2. The majority of interviews were audio recorded and transcribed. In some cases, conditions did not allow for audio recording and field notes were made instead. Interview transcripts were analyzed using line-by-line open coding of transcripts to identify emergent themes.[3]

3 Methods are described in Creswell, J.W. (2013). *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*. London, UK: Sage.

TABLE 2: List of interviews conducted with participating groups*

GROUP	SUBJECTS	DATE
China Group 1	Executive Director, Program Manager (technical projects)	2010
China Group 3	System Administrator	2011
Rights Group 1	Chief Technical Officer, Program Manager	2012, 2014
Rights Group 2	Technical Officer	2011, 2014
Tibet Group 1	Executive Director, Program Director, Program Director (technical projects), Program Officer, Security Trainer	2011, 2012, 2013
Tibet Group 2	Executive Director	2013
Tibet Group 3	Editor-in-Chief	2014
Tibet Group 4	Technical Volunteer	2013
Tibet Group 5	Program Officer	2014

**NOTE: We were unable to conduct a site visit and interview with China Group 2, because they did not maintain participation in the project.*

DATA ANALYSIS

Malware Analysis: We examined malware samples using static and dynamic analysis tools (e.g., IDA and OllyDbg), as well as manual analysis to extract information on exploits, malware functionality, malware family, command-and-control (C2) infrastructure, and other properties of the malware code (e.g., mutex and exported function names).

Email Content Analysis: We reviewed the subject line, body, and attachments for each submitted email and grouped the content into specific themes and categories. The header of each email was analyzed to determine if the sending email address was spoofed or the email address was otherwise designed to appear to come from a real person and / or organization. Indicators drawn from this analysis were used to assess the relative sophistication of the social engineering tactics found in the messages (we incorporate these indicators into our Targeted Threat Index described below). We conducted regular inter-rater reliability checks that flagged any potential edge cases and inconsistencies for discussion and re-evaluation.

Targeted Threat Index: We developed the Targeted Threat Index (TTI), which is a metric that characterizes and quantifies the sophistication of targeted attacks, to provide a consistent ranking of how advanced any given targeted malware attack is. 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.

Cluster Analysis: Through identification of patterns in malware families, development cycles, shared infrastructure, and social engineering tactics, we identified relationships between attacks and, when possible, linked them to known malware campaigns and threat actors.

DATA OVERVIEW
A high level overview of our datasets

EMAIL SUBMISSIONS

The malicious emails analyzed in this report span more than four years, from October 10, 2009 to December 31, 2013. During this period we collected 817 emails from the 10 groups participating in our study.

FIGURE 1: Cumulative number of email submissions per month during the study

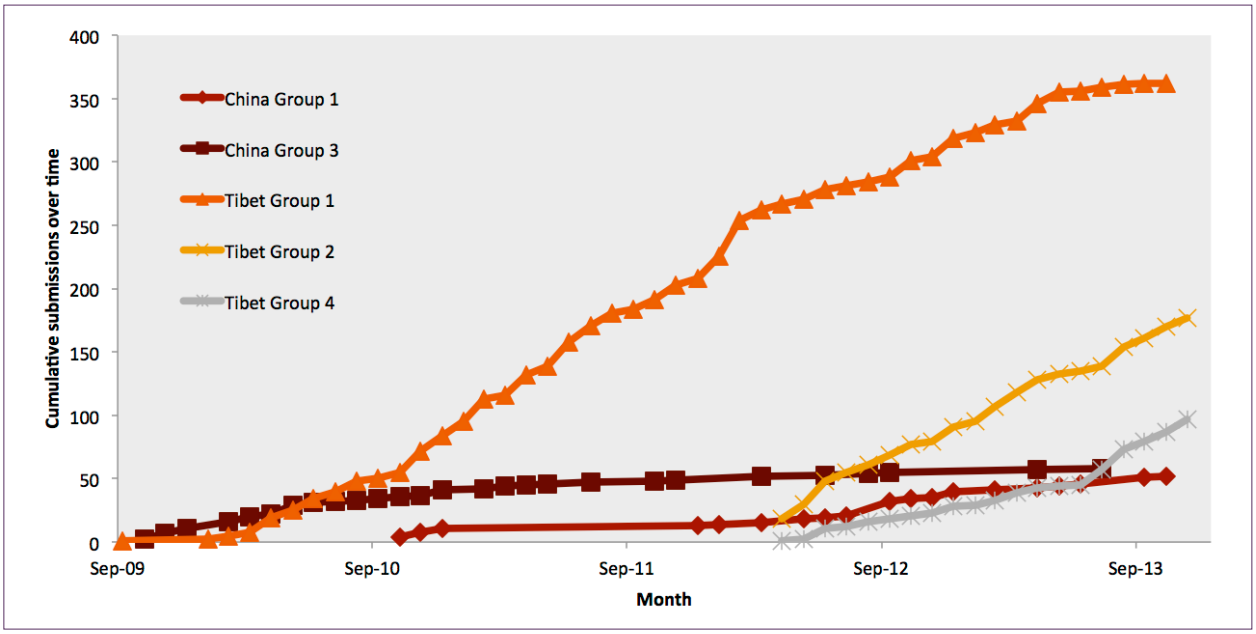
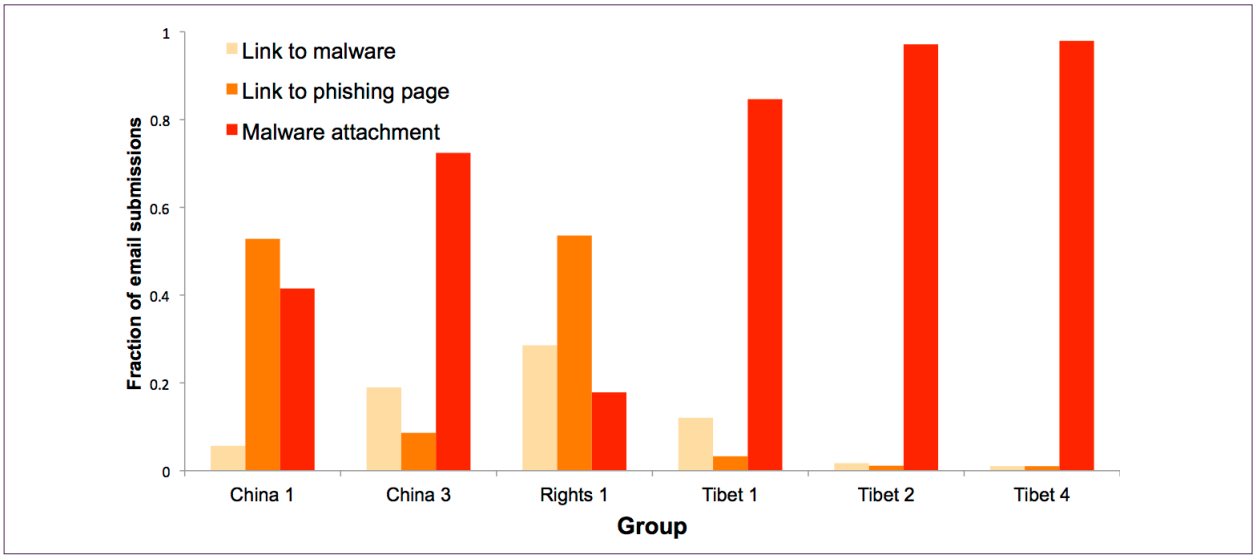


Figure 1 shows the cumulative number of email submissions per month over the course of the study. Although the first formal submission was received on November 28, 2011, some groups had existing archives of malicious messages received by their members, and they provided us with these older emails. Tibet Group 1 ac-

counts for the highest number of submissions relative to the other groups as it was one of the first groups in the study and is persistently targeted. Tibet Groups 2 and 4, which joined the study at a later date (April 2012), show a similar submission rate to Tibet Group 1, suggesting these groups are targeted at a comparable level.

FIGURE 2: Malicious emails by type for groups submitting 25 or more emails



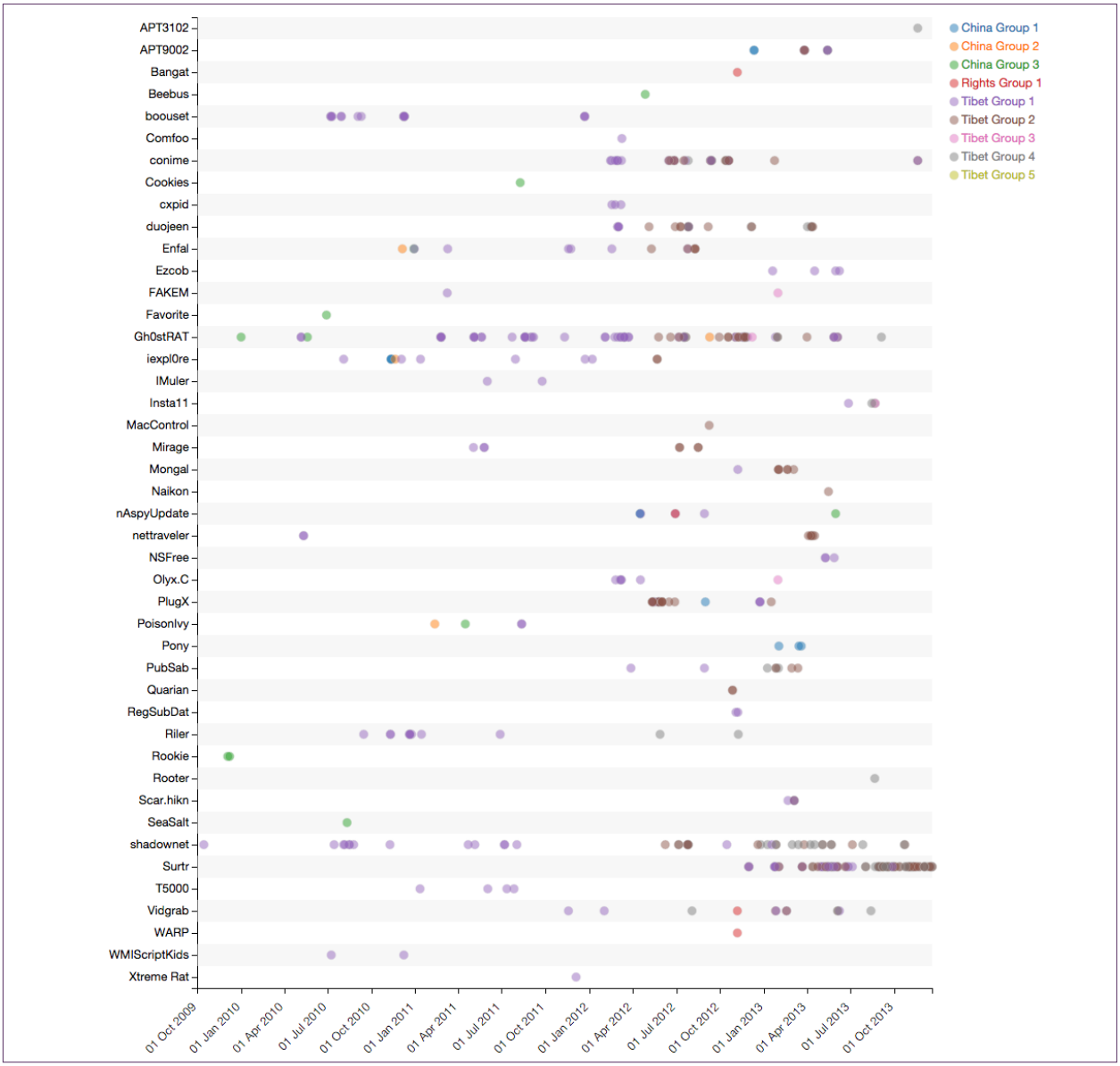
We classify emails as malicious if they include attached malware, a direct link to malware or a drive-by download site, or a link to a phishing page. Figure 2 shows the number of emails of each type for the groups that submitted at least 25 emails to our system. The most common technique employed in these emails was a malicious attachment to the message. However, we observe a higher rate of phishing attacks on the China Groups and the Rights Groups. In particular, 46% of the emails submitted by China Group 1, and 50% of the emails submitted by Rights Group 1, direct the user to a phishing website.

The rate of submissions to our project meant that it was feasible to manually analyze email attachments for malware as they were submitted. This analysis gives us higher confidence in our results than if we had automated the process. Antivirus (AV) signatures frequently fail to detect new or modified threats, and can overlook the kind of malicious payloads that can be identified with manual inspection (e.g., shellcode in an RTF exploit). In total, we analyzed 3,617 payload files and found 2,814 (78%) to be malicious.

MALWARE FAMILIES

We identified malware families through patterns in network traffic and characteristics in the code, such as strings seen in the binaries or names and locations of dropped files. In total, we identified 44 separate malware families (not including variants). The most frequently occurring families are Gh0st RAT, Surtr, Shadownet, Conime, Duojeen, and PlugX.

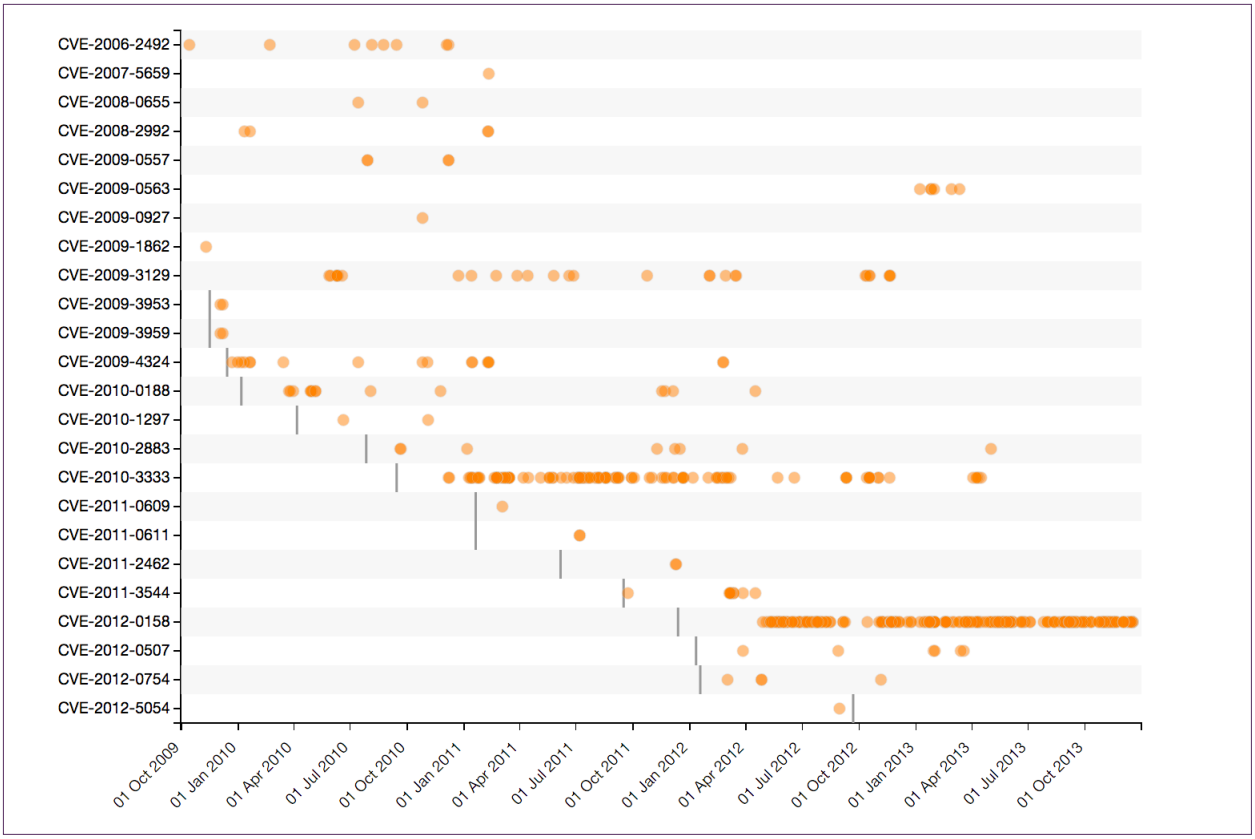
FIGURE 3: Malware family timeline
(The coloured dots represent attacks using a particular malware family against one of our study groups.)



CVEs

Common Vulnerabilities and Exposures (CVEs) is a dictionary of common names for publicly known security vulnerabilities. CVEs are each assigned a unique identifier code, with the form CVE-YYYY-NNNN, where YYYY indicates the year they were identified and NNNN are arbitrary digits. We identified 24 distinct CVEs used in 483 of the email attacks as displayed in Figure 4.

FIGURE 4: CVEs identified in samples during the study period.
(Vertical gray bars represent the date the CVE was created and orange dots represent targeted attacks using that CVE.)



The most common CVEs we observed were CVE-2010-3333 (used in 112 attacks) and CVE-2012-0158 (used in 294 attacks), which are both vulnerabilities in the way Microsoft Word handles RTF documents. Figure 4 clearly illustrates the shift in use from CVE-2010-3333 to CVE-2012-0158 in March and April of 2012. The popularity of these vulnerabilities is not limited to our dataset. They have been widely used in other attacks against a variety of targets.

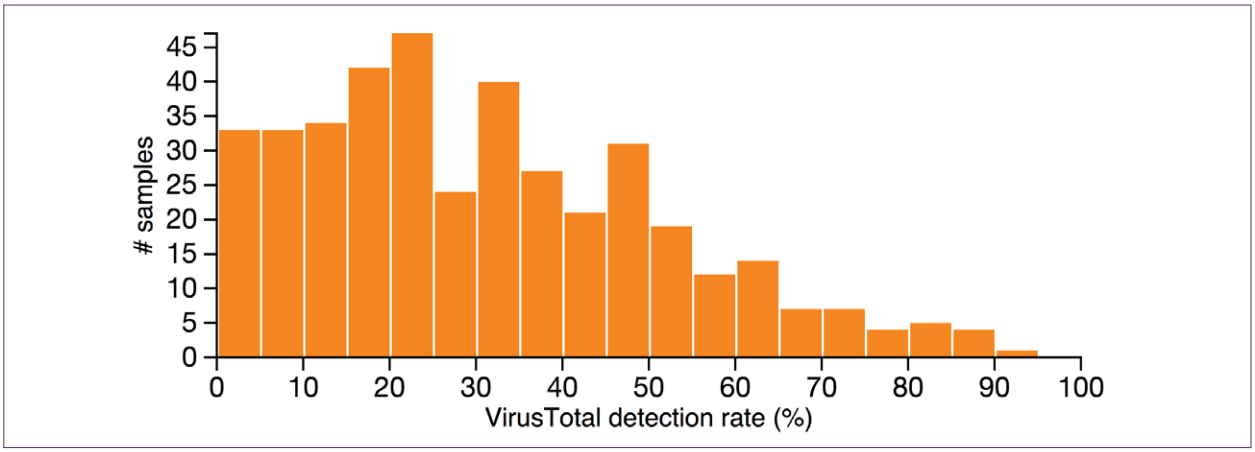
During four years of tracking attacks against our groups, we observed only one zero-day exploit. This attack used the Flash vulnerability CVE-2012-5054, and was sent 22 days before the CVE entry was created.

These results show that vulnerabilities exploited by targeted malware attacks against CSOs are typically not technically advanced (compared with financially-motivated malware and commercial lawful intercept kits), and often use old (patched) vulnerabilities. For example, CVE-2012-0158 has been patched since April 10, 2012, but has remained the most common vulnerability used in attacks against the Tibet Groups for well over a year after the fix was issued. The repeated use of this vulnerability suggests the attackers are achieving successful compromises because target systems did not have the latest security updates. A possible explanation is that licensed software is cost-prohibitive for many organizations in the developing world, while pirated copies are easily available, leading many to use pirated operating systems and software.

ANTIVIRUS DETECTION

VirusTotal is a service that scans files through 53 different AV engines and provides a summary of malware detection results. We find that 369 of the 659 samples we received (56%) had been submitted to VirusTotal at the time of writing, with a median AV detection rate of 24% and mean detection rate of 25%. Detection rates were generally low, as 86% of these samples had a detection rate below 50%, meaning that less than half of the AV packages tested were able to identify them as malicious. These results suggest that simply running AV software, although potentially helpful, is not a very effective defence against these attacks.

FIGURE 5: Histogram of antivirus detection rates provided by VirusTotal



This low detection rate we observed is due in part to the extensive presence of CVE-2012-0158, which uses a number of techniques to hide the vulnerability from AV scanners.

One of the simplest of these detection-reducing techniques is modifying the RTF header, since Microsoft Word will still be able to open the file, but fewer AV scanners will detect it as malicious. Another basic technique is encrypting malicious document and providing a password to open the file in the associated email. Simply adding a password to malicious files can help prevent AV detection.

Since there are four ActiveX controllers—ListView, ListView2, TreeView, and TreeView2—affected by this vulnerability and there are no strict syntax restrictions, there can be a large variance in the document templates into which malicious payloads are inserted. These can cause newer templates to initially have lower detection rates.

A notable technique observed was the createion of a MIME HTML (MHTML) file that uses the vulnerable ActiveX controllers. By default, MHTML files are opened by a browser: however, they can also be opened by Microsoft Word, which will trigger the exploit. Since Microsoft Word may not be the default application to open the file, automated sandbox programs may fail to detect the file as malicious.

The older CVE-2010-3333 vulnerability had similar issues with AV detection, because of the wide number of ways to encode the vulnerability. A small change in the way the vulnerability was written could evade signature detection while remaining functionally the same.

Although AV definitions are updated to account for evasion tricks, the lag between the use of evasion techniques in the wild and definition updates results in temporarily low detection rates, and hence the likelihood of successful compromises.

EMAIL CONTENT ANALYSIS

Subject line, body, and attachments: The content of the subject line, body, and attachments for each submitted email were content coded into 134 categories grouped under eight themes:

- 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 people)
- Political (reference to specific political issues)
- Technology (reference to technical support)
- Miscellaneous (content without clear context or categories that did not fall into one of the other themes)

Email headers: The header of each email was analyzed to determine if the sending email address was spoofed, or the email address was otherwise designed to appear to come from a real person and / or organization (for example, by registering an email account that resembles a legitimate sender's name from a free email provider). We divide the results based on whether they attempted to spoof an organization or a specific person.

Results of this analysis confirm that message content and fraudulent senders are tailored to the interests of the target organizations.

Of the 520 total emails received by the Tibet Groups, 97% referenced content related to Tibetan issues. Email lures leveraged specific events of interest and respected persons in the Tibetan community. Emails referenced Tibet-related events, including holidays (Tibetan New Year), anniversaries (His Holiness the Dalai Lama's birthday), and protests (see Table 3). The most frequently referenced events were Tibetan self-immolations (31% of the emails leveraging event-related content).

Some of the attachments actually cannot be detected as a virus...We're not even sure if it...will cause any harm at all. It's just that the antivirus [is] saying that 'there's no threat,' but obviously there's something wrong with it."

—China Group 1

TABLE 3: Breakdown of top five categories in the Event theme for Tibet Groups

CATEGORY	NUMBER OF EMAIL RECORDS
Self-Immolation *	56
Tibetan National Uprising Day	24
HHDL Birthday	19
Flame of Truth Rally	13
Kalon Tripa Election	9

** Self-immolations are a controversial form of protest that Tibetans have used as a statement of opposition to Chinese government practices concerning Tibet. These protests have escalated in recent years. At the time of writing, it is estimated that since 2009, approximately 132 Tibetans have self-immolated.*

Of the 520 emails received by Tibet Groups, 272 (52%) were designed to appear to come from real organizations. In total 58 organizations were spoofed, of which 53 (91%) were Tibet-related groups (see Table 4). The most frequently spoofed organization was the Central Tibetan Administration. The identities of four of the Tibet Groups in our study (Tibet Groups 1, 2, 3, and 5) were frequently spoofed internally and to external contacts. The frequency of emails with fraudulent contacts from Tibetan organizations shows an effort to have the message appear to come from within the Tibetan community and leverage existing trust relationships.

> *The emotions of the immolations [are] being used against people to have them click on [attachments].*"
>
> —Tibet Group 1

TABLE 4: Breakdown of top five categories in the Spoofed Organizations theme for Tibet Groups

CATEGORY	NUMBER OF EMAIL RECORDS
Central Tibetan Administration	58
Tibet Group 1	26
Tibet Group 2	13
Tibet Group 5	13
Tibet Group 3	11

We see a similar pattern for the China Groups. Of the 48 emails received by the China Groups, 46 (95%) referenced China. Content included references to Chinese political events such as the Communist Party of China (CPC) 18th Party Congress; the June 4, 1989 Tiananmen Square crackdown; and Chinese dissidents and prominent members of the CPC (see Table 5). Of the 48 emails, 13 (27%) spoofed real organizations (see Table 6). Two of our China Groups were spoofed (China Group 1, China Group 3). Rights Group 1 was also spoofed in one message to China Group 1. The remaining spoofed organizations were prominent human rights groups and intergovernmental organizations (e.g., the UN Office of the High Commissioner for Human Rights).

TABLE 5: Breakdown of top five categories in the Event theme for China Groups

CATEGORY	NO. OF EMAIL RECORDS
Jasmine Revolution	8
June 4, 1989, Tiananmen Square Crackdown	4
CPC 18th Party Congress	2
July 2009 Urumqi Riots	1
Chinese New Year	1

TABLE 6: Breakdown of top five spoofed organizations for China groups

CATEGORY	NO. OF EMAIL RECORDS
China Group 1	4
China Group 3	3
Office of the High Commissioner for Human Rights	3
Open Society Institute	2
Chinese Human Rights Defenders	2

The volume of email submissions from Rights Group 1 and Rights Group 2 was much lower than that from the Tibet and China Groups. However, we also observed content and email senders tailored to these organizations. Rights Group 1 received messages related to human rights issues in Africa and Russia. Of the 12 emails submitted, 92% were made to appear to come from Rights Group 1 email addresses (no other organizations were spoofed). The majority of these messages were phishing attempts with lures related to IT support, designed to gain access to Rights Group 1 email credentials. Rights Group 2 submitted two email samples, both of which were related to human rights issues in the Middle East. One message was made to appear to come from a Rights Group 2 email address.

While the content analysis results clearly show targeted attacks tailored to the interests of targeted groups, content coding alone does no provide a measure of the sophistication of social engineering used in the attacks. In the following section, we describe a metric to determine relative sophistication of attacks.

TARGETED THREAT INDEX

Our dataset includes a wide range of targeted malware threats that have varying levels of complexity. This range presents a challenge in ranking the relative sophistication of the malware and targeting tactics used by attackers.

While metrics such as the Common Vulnerability Scoring System 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 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 deeper analysis of incoming threats, as well as for getting an overall idea of how severely an organization is threatened.[4]

The TTI Score is calculated in two parts: (Social Engineering Sophistication Base Value) × (Technical Sophistication Multiplier) = TTI Score

TTI scores range from zero to 10, where 10 is the most sophisticated attack. Scores of zero are reserved for threats that are not targeted, even if they are malicious. For example, an email from a widely-spread spam campaign using an attached PDF or XLS file to bypass anti-spam filters would score zero. Sophisticated financially-motivated malware would also score zero if it was not part of a *targeted* attack.

Social Engineering Sophistication

To measure the targeting sophistication base value we assign a score that ranges from zero to five, which rates the social engineering techniques used to persuade a victim to open a malicious link or attachment. This score considers the content, presentation, and claimed sender identity of the email. 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 features for each score are detailed in Table 7(for examples of emails with each of these scores see Appendix A).

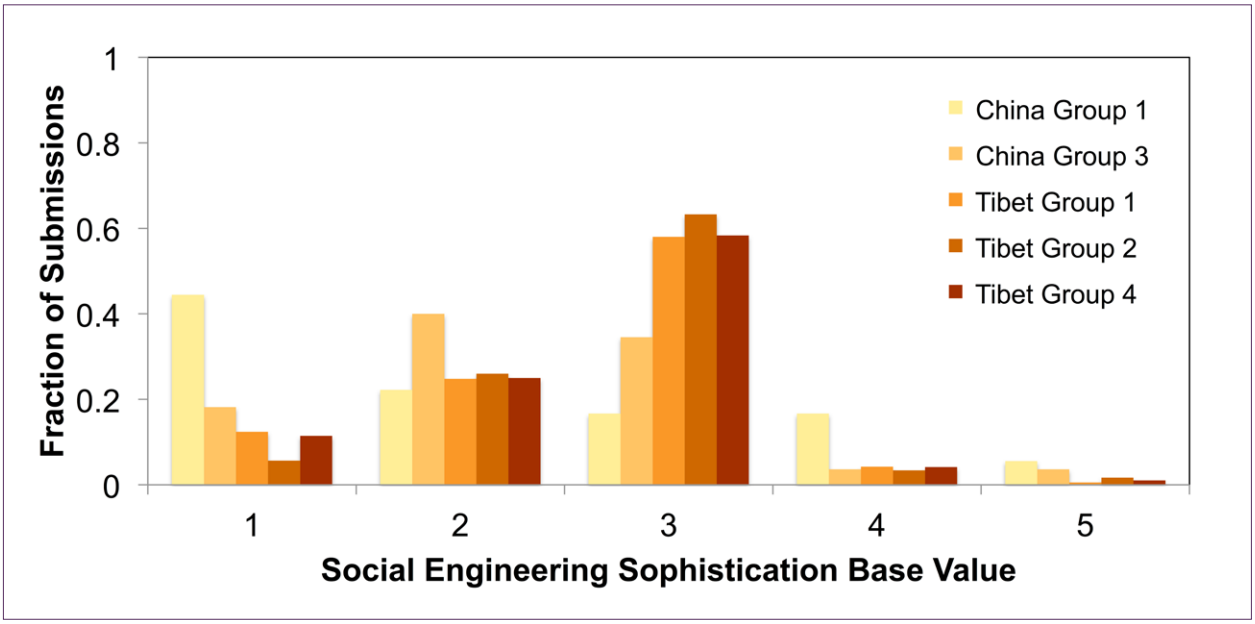
4 For further details on the TTI including detailed discussion of its design, limitations, and plans for future work see: Hardy, S., Crete-Nishihata, M., Kleemola, K., Senft, A., Sonne, S., Wiseman, G., Gill, P., Deibert, R. "Targeted Threat Index: Characterizing and Quantifying Politically-Motivated Targeted Malware." USENIX Security 2014. https://www.usenix.org/system/files/conference/usenixsecurity14/sec14-paper-hardy.pdf

TABLE 7: TTI base value score

VALUE	DESCRIPTION
0	**Not targeted** . Recipient does not appear to be a specific target. . Content is not relevant to recipient. . The email is likely spam or a non-targeted phishing attempt.
1	**Targeted, Not customized** . Recipient is a specific target. . Content is not relevant to recipient or contains information that is obviously false with little to no validation required by the recipient. . The email 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 email 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, or a conference or event website) and can be externally verified (e.g., message references information that can be found online). Or, the email text appears to repurpose legitimate email messages that may have been collected from public mailing lists or from compromised accounts. . The email header and / or signature references a real person or organization.
4	**Targeted, Personalized** . Recipient is a specific target. . Email 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 email header and / or signature references a real person or organization.
5	**Targeted, Highly personalized** . Recipient is a specific target. . Email 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 email header and / or signature references a real person or organization.

Figure 6 shows the targeting score for organizations in our study that submitted at least 50 emails. We can see that attackers targeting these groups put significant effort into the message lures. In particular more than half of the messages targeting the Tibet Groups in Figure 6 have a targeting score of 3 or higher. This result means threat actors are taking care to make the email 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, were rare, but we nevertheless observed cases. For example, in the case of China Group 3, we observed an email that claimed to be from one of the organization's funders and referenced a specific meeting they had planned that was not public knowledge (social engineering score: 5).

FIGURE 6 : Social engineering base value of emails submitted per group (minimum 50 submissions)



Technical Sophistication

The technical sophistication multiplier ranks the relative technical sophistication of malware. This score is determined by measuring how well the payload of the malware conceals its presence on a compromised machine. We use a multiplier because advanced malware requires significantly more resources 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:

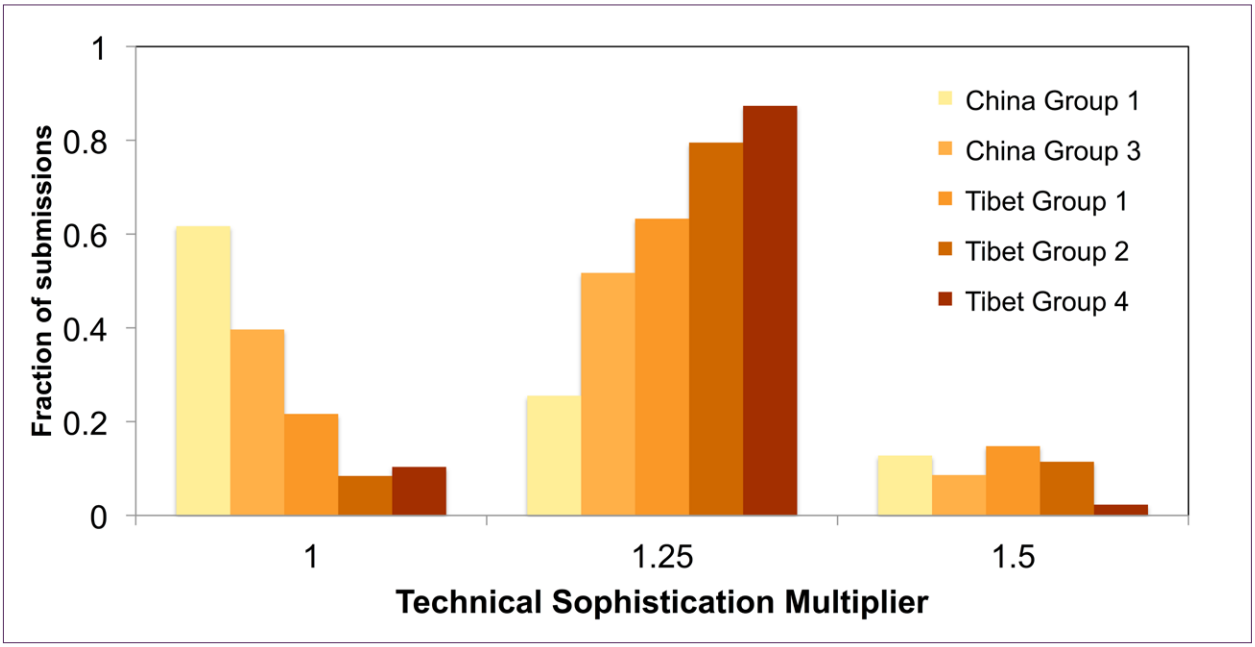
- It allows the compromised system to remain infected for a longer period;
- It hinders analysts from dissecting a sample, developing instructions to detect the malware, and disinfecting a compromised system; and
- Since most commonly used remote access trojans (RATs) have the same core functionality (e.g., key-logging, 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.

TABLE 8: TTI technical sophistication multiplier

VALUE	DESCRIPTION
1	**Not protected** The sample contains no code protection, like 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, including: code protection using publicly available tools where the reversing method is available (e.g., UPX packing); simple anti-reversing techniques like not using import tables, or a call to IsDebuggerPresent(); self-disabling in the presence of antivirus software.
1.5	**Multiple minor protection techniques** The sample contains multiple distinct minor code protection techniques (anti-reversing tricks, packing, virtual machine / 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.

Figure 7 shows the technical sophistication multiplier values for emails submitted by the different organizations in our study. Our results show that malware used to target the groups in our study was relatively simple. The highest multiplier value we observed 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 Tibet Groups observing a higher fraction of malware rated 1.25 and China Groups observing a higher fraction rated 1.

FIGURE 7: Technical sophistication multiplier of emails submitted per group (minimum 50 submissions)



Targeted Threat Index Results Overview

The TTI metric can help us better characterize the relative threat posed by targeted malware in several ways. Table 9 shows fthe technical sophistication multiplier and maximum / minimum TTI scores for malware families observed in our dataset. Since we primarily find 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 threats are in practice.

If we consider the malware families with the highest technical sophistication, we can

see that their TTI values are relatively low, with scores mostly ranging from 1.5 to 4.5 (and one notable exception of 7.5). These tend to be malware families that are regularly used in targeted malware campaigns known to researchers. In particular, PlugX and PoisonIvy have been found used together in targeted attacks, and PlugX is still in active use and under continuous improvement. Despite their technical sophistication, these threats are not well executed and pose less of a risk to CSOs in which users may be able to identify and avoid these threats.

In contrast, the top five malware families in terms of TTI have lower technical sophistication multipliers (1.25) but much higher levels of social engineering. A notable exception is one highly targeted attack (social engineering score 5.0) that used PlugX (technical sophistication score 1.5) resulting in a TTI value of 7.5 (the highest score in the dataset). While this attack has a higher technical sophistication score than the top five malware families, the high TTI score is due to the level of targeting.

TABLE 9: Top malware families in our dataset by technical sophistication multiplier and final TTI score

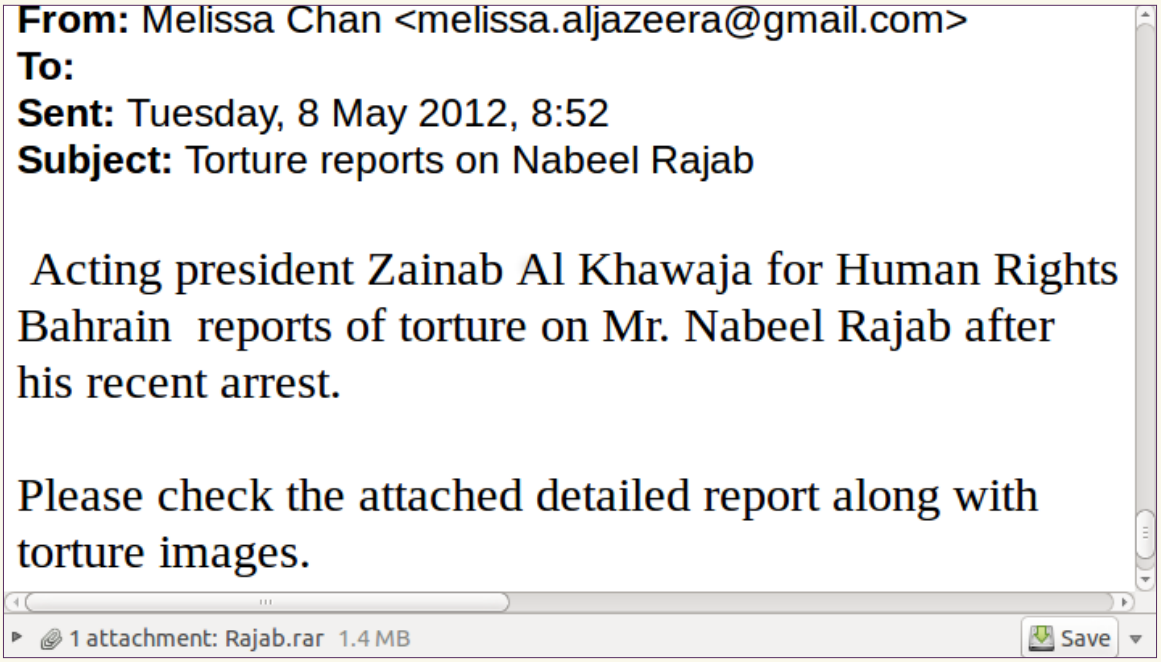
TECHNICAL SOPHISTICATION		
Family	**Max TTI**	**Technical Sophistication**
PlugX	7.5	1.5
GhOst RAT (LURKO), ShadowNet	6.25	1.25
Conime, Duojeen, IEXPLORE, GLASSES, cxpid, Enfal, Surtr, Vidgrab	5	1.25
Cookies	5	1.0

TTI		
Family	**Max TTI**	**Technical Sophistication**
3102	3	1.5
nAspyUpdate	1.5	1.5
PlugX	7.5	1.5
PosionIvy	3	1.5
WMIScriptKids	3	1.5

ANALYZING COMMERCIAL SPYWARE WITH THE TTI

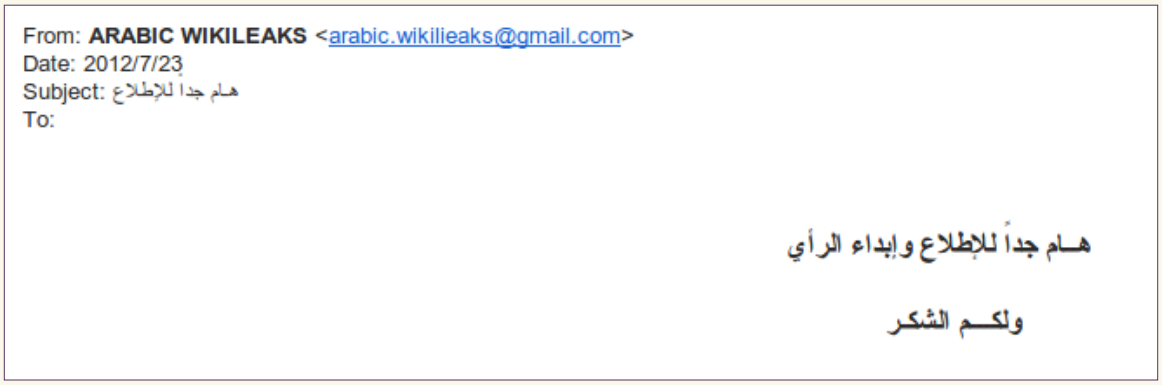
Attacks using advanced commercial spyware such as FinFisher and DaVinci RCS do not necessarily rank higher on the TTI.

We analyzed a sample of FinFisher used against Bahraini activists and evaluated it with the TTI. The malware sample is technically advanced, scoring a 2.0, as a result of multiple advanced protection techniques, including a custom-written virtualized packer, MBR modification, and rootkit functionality. However, the email used in the attack is poorly customized and has several attributes that made it look suspicious to the intended target. The email attempts to reference an NGO called Bahrain Center for Human Rights, but mistakenly refers to it as "Human Rights Bahrain." The message also lists the wrong name for the acting president of the group. It appears to come from a real journalist, Melissa Chan of Al Jazeera, but provides a suspicious gmail address (melissa.aljazeera@gmail.com). These attributes give the email a social engineering base value of 2. As a result, the attack scores an overall TTI score of 4.0, which is relatively low compared to many other attacks seen in our study. This result shows the importance of social engineering tactics: FinFisher is only effective if it is surreptitiously installed on a user's computer, which in some cases requires opening a malicious file (however, both FinFisher and Hacking Team offer optional network injection products that permit remote attackers to infect a device without user interaction).



Analyzing Commercial Spyware with the TTI (Cont'd)

Similar results can be observed with respect to attacks using DaVinci RCS, developed by Italy-based company Hacking Team, which has been used against activists and independent media groups. RCS also scores a 2.0 on our technical sophistication scale. We analyzed a targeted attack using RCS against a dissident in the United Arab Emirates. The email appears to come from "Arabic Wikileaks" (arabic.wikileaks@gmail.com) and asks the recipient to read a "very important message." Again, while the malware used in these attacks is technically sophisticated, the social engineering lure is poorly customized (social engineering base value 2), resulting in an overall TTI score of 4.0.


From: **ARABIC WIKILEAKS** <arabic.wikilieaks@gmail.com>
Date: 2012/7/23
Subject: هـام جدا للإطلاع
To:

هـام جداً للإطلاع وإبداء الرأي

ولكـم الشكـر

These results suggest that different threat actors possess varying levels and types of resources, and as a result use different attack methods. The majority of malware submitted in our study appears to be from actors that have in-house malware development capabilities, and the capacity to organize targeted campaigns. However, as this report shows, in many cases they spend significant effort on social engineering, but generally do not use technically advanced malware. Conversely, operators of FinFisher and DaVinci RCS have purchased advanced malware products, but in some cases paired them with relatively unsophisticated social engineering.