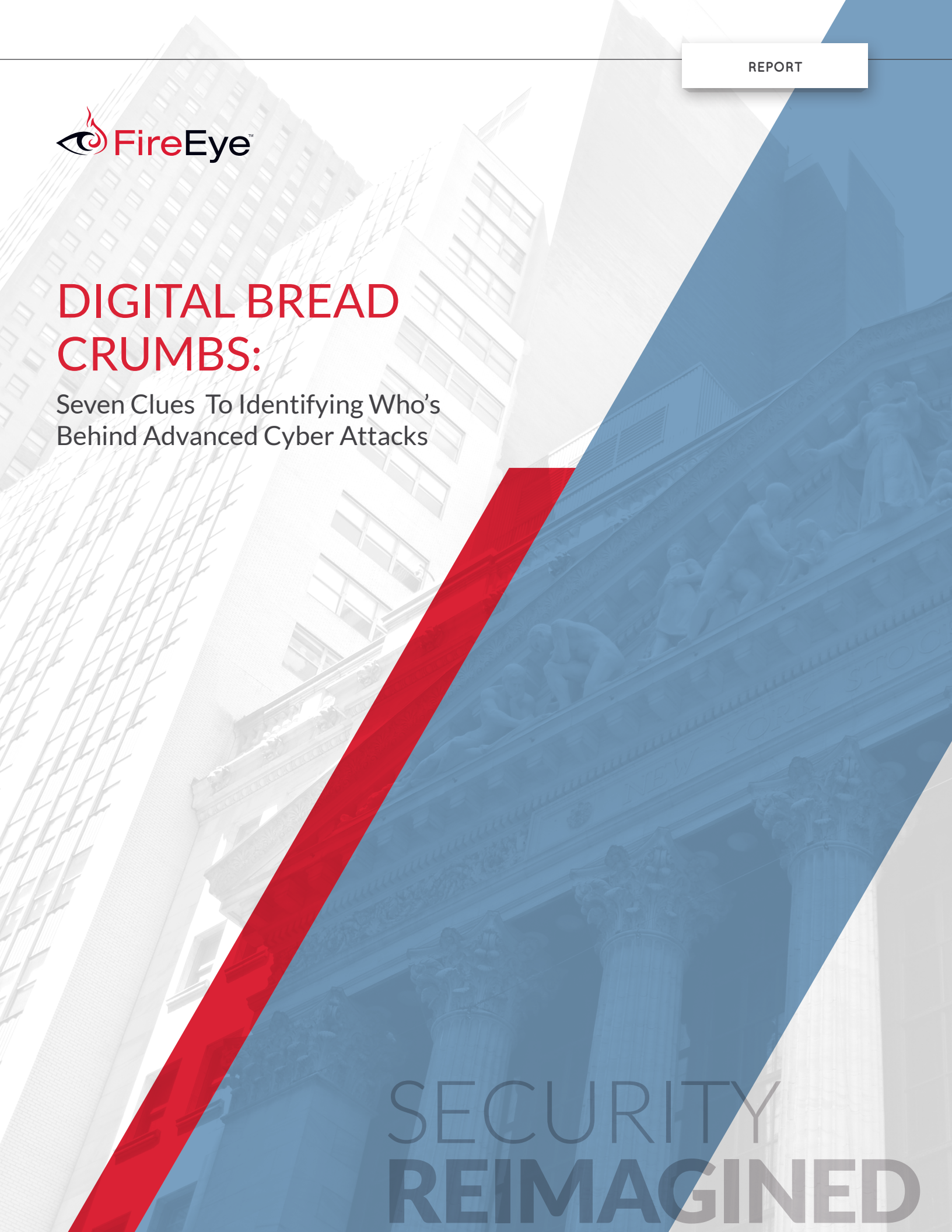
REPORT

FireEye

DIGITAL BREAD CRUMBS:

Seven Clues To Identifying Who's Behind Advanced Cyber Attacks



SECURITY REIMAGINED

CONTENTS

Executive Summary

In today's cyber threat landscape, identifying your enemy is a crucial piece of any defense plan. Finding out who your attackers are, how they work, and what they want is critical to protecting your data and intellectual property.

Fortunately, breached computer systems, like any crime scene, contain a trail of clues. When it comes to advanced cyber attacks, attackers may give themselves away inside their malware code, phishing emails, command-and-control (CnC) servers used, and even behavior. Just as the science of fingerprints, DNA, and fiber analysis have become invaluable in criminal forensics, connecting the dots of an advanced cyber attack can help identify even the most sophisticated threat actors—if researchers know what to look for.

Drawing from a sample of nearly 1,500 campaigns tracked by FireEye®, this paper describes the following facets of malware attacks and what they often reveal about the culprits:

- **Keyboard Layout.** Hidden in phishing attempts is information about the attacker's choice of keyboard, which varies by language and region.
- **Malware Metadata.** Malware source code contains technical details that suggest the attacker's language, location, and ties to other campaigns.
- **Embedded Fonts.** The fonts used in phishing emails point to the origin of the attack. This is true even when the fonts are not normally used in the attacker's native language.
- **DNS Registration.** Domains used in attacks pinpoint the attacker's location. Duplicate registration information can tie multiple domains to a common culprit.
- **Language.** Language artifacts embedded in malware often point to the attacker's country of origin. And common language mistakes in phishing emails can sometimes be reverse-engineered to determine the writer's native language.
- **Remote Administration Tool Configuration.** Popular malware-creation tools include a bevy of configuration options. These options are often unique to the attacker using the tool, allowing researchers to tie disparate attacks to a common threat actor.
- **Behavior.** Behavioral patterns such as methods and targets give away some of the attacker's methods and motives.

By examining these areas, security professionals can make great strides in identifying threat actors and better defend their organizations against future cyber assaults.

Introduction

Although cyber attacks have grown more advanced and tenacious in recent years, there is still no such thing as the perfect crime. Every stage of the attack kill chain—reconnaissance, weaponization, delivery, exploitation, installation, command and control, and actions on objectives (usually exfiltration)[1]—can leave behind a digital paper trail.

That is because every stage requires some point of contact between the attacker and the target. Sometimes, this contact is direct, such as a phishing email. At other times, this contact is indirect, such as during a callback connecting the target computers with the attacker's system. Both types of contact are opportunities to learn more about the attacker. Analyzed correctly, that information can help security professionals better contain the damage, fix breached systems, and anticipate future attacks.

A note of caution: Although the digital forensics techniques spelled out in this report have proven helpful to FireEye researchers, clues are often misleading and contradictory. Analyzing evidence is intricate and painstaking—a delicate mix of science and art that rarely uncovers any single "smoking gun." Cybercriminals are experts at misdirection, so take no sign at face value. Before reaching any conclusions about the source of an attack, FireEye strongly recommends weighing evidence from multiple sources and enlisting digital forensics experts.

1. Keyboard Layout

Researchers can determine the layout of the keyboard used to create a given piece of malware by examining the "charset" attribute of the email header in phishing emails. Most phishing attempts use standard keyboard layouts that do not point to any particular country. But when a nonstandard keyboard is apparent, it is a strong indicator.

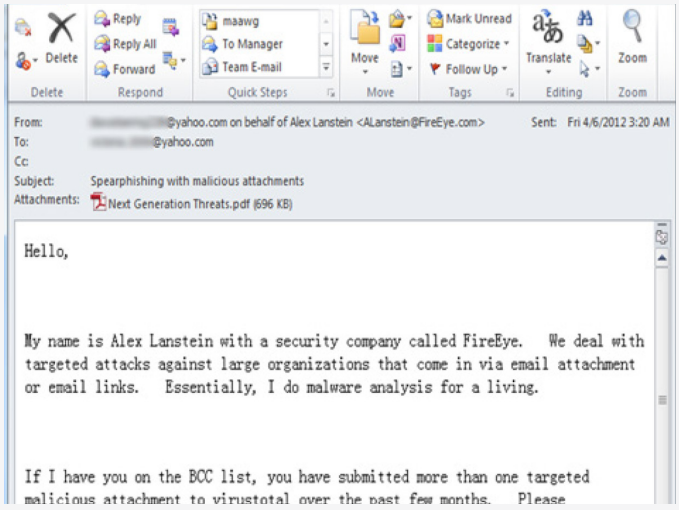
FireEye researchers have found that many aspects of malware campaigns have the earmarks of being typed on a Mandarin (GB2312) keyboard used in China. In a similar vein, North Korea's KPS 9566 character set can help identify the campaigns that emanate from that region.

This method of tracing the origins of an attack is not foolproof. In theory, a Russian national could employ a North Korean keyboard to disguise his or her identity and whereabouts, for example.

In March 2012, FireEye researcher Alex Lanstein emailed a number of Tibetan activists to warn them that they were targets of a cyber attack. The attackers subsequently obtained a copy of Lanstein's email from one of the targets and used it to bait other activists. Unlike the original email, which used a standard Western-language keyboard (Windows-1252), the decoy came from a sender who used China's GB2312 keyboard layout.

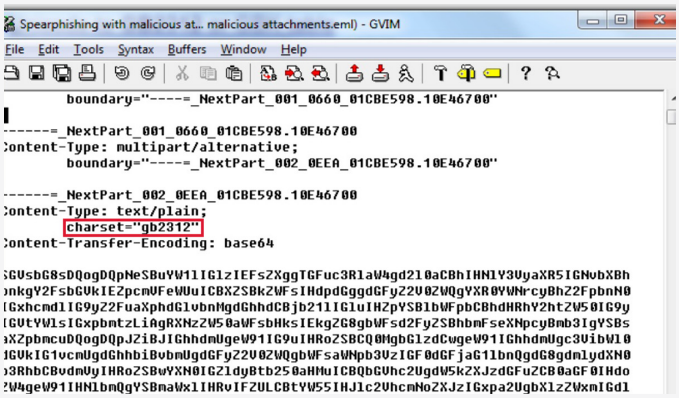
Figure 1 shows the decoy email. Figure 2 shows the email header information that reveals the keyboard layout.

Figure 1:
Decoy phishing email sent to Tibetan activists



[1] Eric M. Hutchins, Michael J. Cloppert, and Rohan M. Amin (Lockheed Martin). "Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains." November 2010.

Figure 2:
Charset coding in phishing email (see Figure 1)



2. Malware Metadata

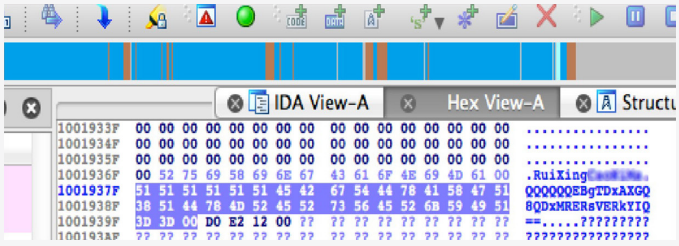
Malware's executable code often references the original source directory that organizes source code. In the same way, programs written in C++ reference a project name.

This underlying code can reveal the attacker's language or country of origin, even when the code and other aspects of the attack are tailored to the language of the target.

Figure 3 shows the source code of a recent third-stage attack. Here, the attacker insults Chinese anti-virus software maker Beijing Rising (spelled phonetically as "Ruixing") International Software Co. with a vulgar term.

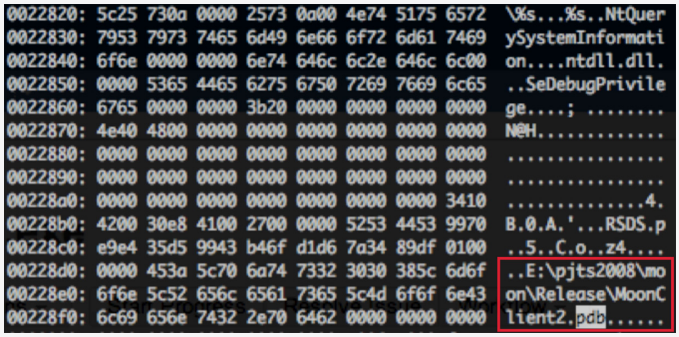
Figure 4 shows the source code of a previously unpublished second-stage attack, an executable file disguised as a PNG file. It was delivered to the endpoint after the initial compromise. The code includes a reference to a process-debugging (PDB) file in the malware writer's hard drive at "E:\pjts2008\moon\Release\MoonClient2.pdb". (PDB files are created for programs written in the Windows .Net framework.) The "MoonClient" file referenced here is a variant of WEBC2 malware used by the Chinese hacker group APT1, also known as the CommentGroup.

Figure 3:
Malware source code insulting Chinese anti-virus software maker Beijing Rising (spelled phonetically as "Ruixing").
We have blurred the insult in this screenshot.



FireEye

Figure 4:
Decoded executable file (PDB reference highlighted)

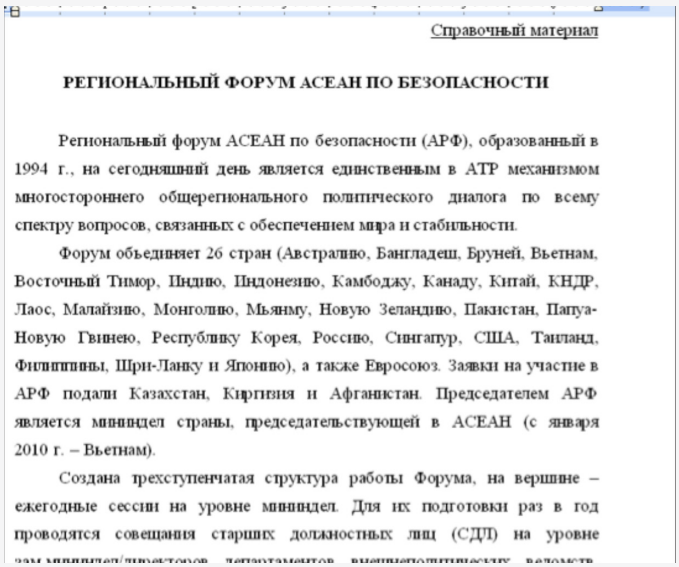


3. Embedded Fonts

Like the charset attribute described in the "Keyboard" section, the font used in phishing emails and other malicious documents can sometimes be useful for tracking back to the source of an APT attack.

Consider the example of the Sanny APT that FireEye researchers discovered recently. Figure 5 shows the decoy document used to lure targets.

Figure 5:
Decoy document written in Russian characters but using a Korean font



Справочный материал

РЕГИОНАЛЬНЫЙ ФОРУМ АСЕАН ПО БЕЗОПАСНОСТИ

Региональный форум АСЕАН по безопасности (АРФ), образованный в 1994 г., на сегодняшний день является единственным в АТР механизмом многостороннего общерегионального политического диалога по всему спектру вопросов, связанных с обеспечением мира и стабильности.

Форум объединяет 26 стран (Австралию, Бангладеш, Бруней, Вьетнам, Восточный Тимор, Индию, Индонезию, Камбоджу, Канаду, Китай, КНДР, Лаос, Малайзию, Монголию, Мьянму, Новую Зеландию, Пакистан, Папуа-Новую Гвинею, Республику Корея, Россию, Сингапур, США, Таиланд, Филиппины, Шри-Ланку и Японию), а также Евросоюз. Заявки на участие в АРФ подали Казахстан, Киргизия и Афганистан. Председателем АРФ является министр страны, председательствующей в АСЕАН (с января 2010 г. – Вьетнам).

Создана трехступенчатая структура работы Форума, на вершине – ежегодные сессии на уровне министр. Для их подготовки раз в год проводятся совещания старших должностных лиц (СДЛ) на уровне

Although the decoy document was written in Russian to target Russian interests, it used the Korean fonts Batang and KPCheongPong. Those font choices reconfirmed existing evidence from other sources that pointed to North Korea, including the author's name and the CnC servers used in the attack. Taken together, the evidence presented a convincing case as to the attacker's origins.

4. DNS Registration

In some cases, threat actors pay to register domains to evade detection by standard malware defenses such as domain blacklists. Often these DNS registrations point directly to an attacker's country of origin.

Even DNS registrations with fake names and addresses can be useful in pinpointing the culprit. In some cases, attackers reuse the bogus contact information across multiple domains. That copying allows researchers to quickly link multiple attacks to a single threat actor and piece together information gleaned from each of the attacks..

Case in point: the so-called "Sin Digoo Affair." Between 2004 and 2011, someone using a Hotmail email address registered several domains under the same names.

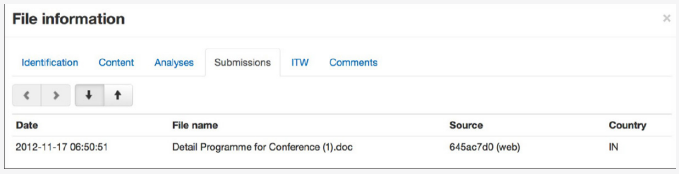
The registrant listed a physical address as a P.O. Box in the town of "Sin Digoo, Californa," apparently a phonetic misspelling of "San Diego." Thanks to the duplicate registration info, researchers were able to connect individual malware attacks to a larger pattern of advanced persistent threats.[2] Korean fonts Batang and KPCheongPong. Those font choices reconfirmed existing evidence from other sources that pointed to North Korea, including the author's name and the CnC servers used in the attack. Taken together, the evidence presented a convincing case as to the attacker's origins.

Similarly, malware researcher Nart Villeneuve used DNS registration information to link China's Zhejiang University to a 2010 attack on Amnesty Hong Kong, journalists, and human-rights activists.[3]

FireEye recently used DNS registration details to link several malware samples uploaded to virus-checking website VirusTotal (see Figure 6). The attacker appears to have uploaded the samples to test whether the anti-virus community was detecting it.

The threat actor who uploaded this sample masks the first-stage CnC attempt. But the second stage, revealed only by running the malware in a live infrastructure, uses the secureplanning.net domain (Figure 7) registered to someone at a presumably fake New Delhi address.

Figure 6:
Example of malware sample upload



[4] Joe Stewart (Dell SecureWorks): „The Sin Digoo Affair", Februar 2012.
[5] Nart Villeneuve: „Nobel Peace Prize, Amnesty HK and Malware", November 2010.

The registration information is not a perfect indicator; a sophisticated attacker could generate deceptive contact information to throw researchers off the trail. But in this case, FireEye researchers observed slightly morphed versions of the malware uploaded to VirusTotal more than 15 times. All of the samples attempted to connect to domains registered with the same New Delhi address, establishing a clearer pattern.

5. Language

Often, many indicators suggest that the language used in a malware campaign is not that of a native speaker. Sometimes those indicators can even point to the attacker's origin.

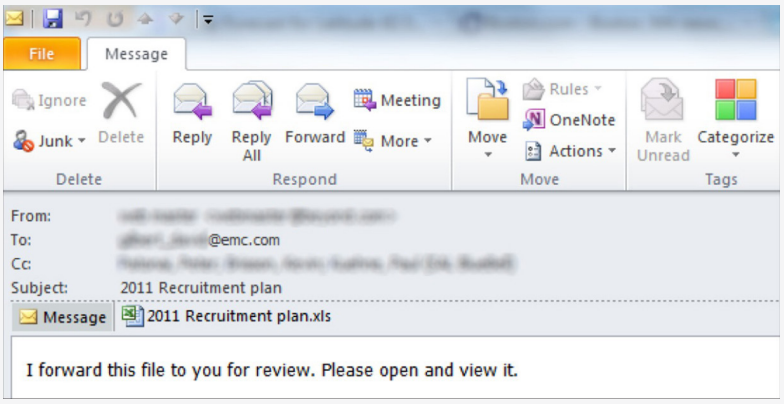
Obvious typos and misspellings are clear signs. In other cases, a more detailed analysis shows telltale signs that the attacker used a language translation site. Knowing how popular translation sites handle certain words and phrases, researchers can determine the original language of phishing emails used in an attack.

Take the much-publicized 2011 attack against RSA. Two groups believed to be working for a government breached the company's network to extract data about RSA's SecurID products. The attack leveraged an unknown vulnerability in Flash, suggesting a high level of technical sophistication. But as shown in Figure 8, the phishing email used poor English and a clumsy (though ultimately successful)

Figure 7: Upload information for sample malware uploaded to VirusTotal

```
POST /download/ad.php HTTP/1.0
Accept: text/plain, text/html
Content-Type: multipart/form-data; boundary=----------------------
User-Agent: Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV2)
Host: secureplanning.net
Content-Length: 4565
Pragma: no-cache
```

Figure 8: Phishing email used in a 2011 attack against RSA



FireEye

appeal to open the attachment. Those characteristics suggest that the individual attacker is not a native English speaker.

Other times, languages embedded inside the malware can help pinpoint the attackers. In Figure 9, a snippet of code for the Backdoor.LV malware uses Arabic names and languages to "tag" targets.

The string of characters decodes to "HacKed By Fayez Hacker_400CD510" as shown in Figure 10.

The code shown in Figure 11 appears to be from the same attacker. The string of characters, when decoded, shows the "400CD510" tag (see Figure 12), this time with Arabic lettering.

Figure 9:
A snippet of the Backdoor. LV malware. The highlighted portion decodes to string shown in Figure 10.

Figure 10:
Backdoor.LV Decoded

Figure11:
Another malware snippet linked to Fayez. Highlighted portion decodes to string shown in Figure 12

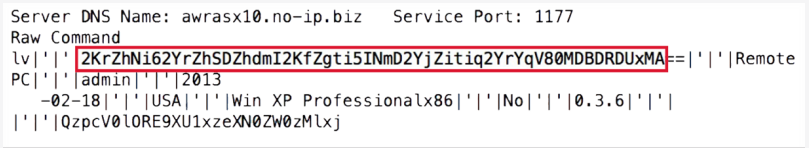
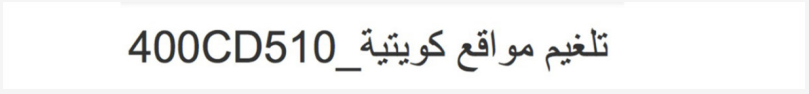


Figure12:
The 400CD510 reference, with Arabic lettering



6. Remote Administration Tool Configuration

Remote Administration Tools (RATs) are a type of malware that give attackers real-time control of a target's computer. The tools support a variety of features such as key logging, screen capturing, video capturing, file transfers, system administration, and command-shell access. Publicly available for a price or even free, RATs appeal to attackers because they are usually well tested and full featured.

Figure 13: Poison Ivy's connection configuration windows (ID and password fields highlighted)

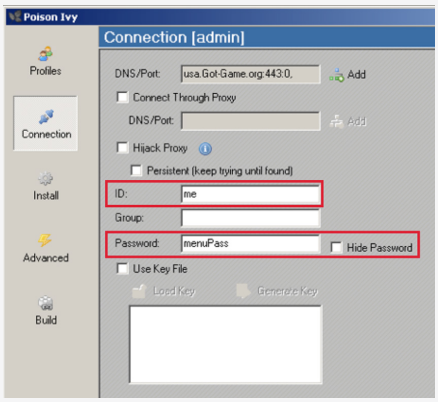
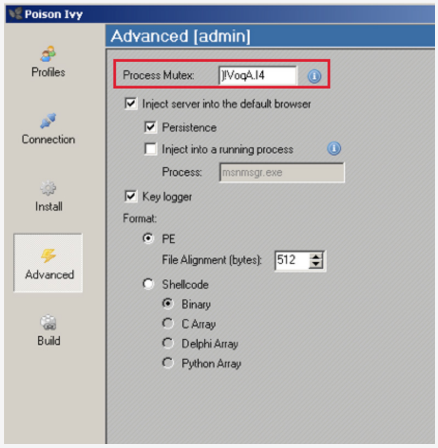


Figure 14: Poison Ivy's advanced configuration window (process mutex field highlighted) highlighted)



RATs might seem to make attribution more difficult; anyone can use them, and many different groups employ the same tools. But their many customization options create a combination of settings distinctive to each attacker. Multiple attacks using a RAT configured in the same way point to a common attacker.

One example is a popular eight-year-old RAT called Poison Ivy. Some of the tool's most revealing configuration options include ID, group, password, and mutex.

Figure 13 shows the ID and password fields in Poison Ivy's connection configuration window. Figure 14 shows the mutex field in the advanced configuration window.

These configuration options can be extracted from compiled RATs using Volatility, an open-source file memory-forensics framework that operates on memory dumps.

In Poison Ivy, the ID and group fields are set by the attacker to tag and organize groups of targets. When the same IDs or group names appear in multiple attacks, researchers can conclude that the attacks are linked.

The password field is used as a key for encrypting Poison Ivy's communications. It is set to "admin" by default and often remains unchanged. But when actively set, passwords can serve as a fingerprint of sorts. They are usually unique and frequently reused by attackers in targeted attack campaigns.

In software, mutex is a program object used to ensure that multiple threads of a program do not attempt to use the same resources at the same time. In Poison Ivy, the mutex serves as a marker to determine whether the tool is already running on an infected system so that

it does not run more than one instance of itself. Poison Ivy uses a default mutex value of)!VoqA.I4. Like passwords set in Poison Ivy, any non-default value is usually unique, which makes it a helpful marker.

7. Behavior

People are creatures of habit. Like anyone, threat actors often demonstrate consistent patterns of behavior over time. They focus on the same targets, use the same CnC servers, and focus on the same industries. These repeated tactics can reveal the approaches, objectives, and whereabouts

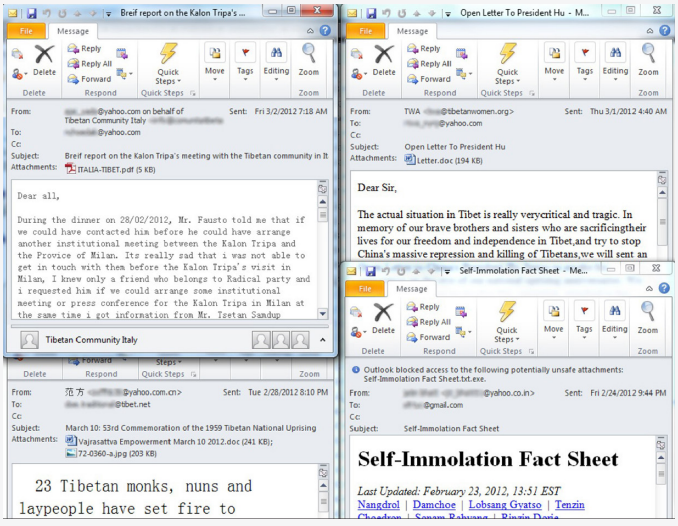
of attackers. That is where threat-actor profiling can help. Much like criminal profiling helps detectives focus on potential suspects, security professionals can observe attackers over time and note patterns. Using that information, researchers can spot the proclivity of a given group toward certain styles and approaches.

In the same way, the attacker's exploit toolkits and tactics also help profile the attacker. Figure 15 shows four separate attacks that use different exploits, different lures, and different first-stage malware implants. But they all target religious activists. And as the heading information reveals (see Figure 16), they are all sent from the same server—some by the Yahoo! Web email service and some by means of a script. This evidence points to multiple actors on the same team, using the same infrastructure.

Conclusion

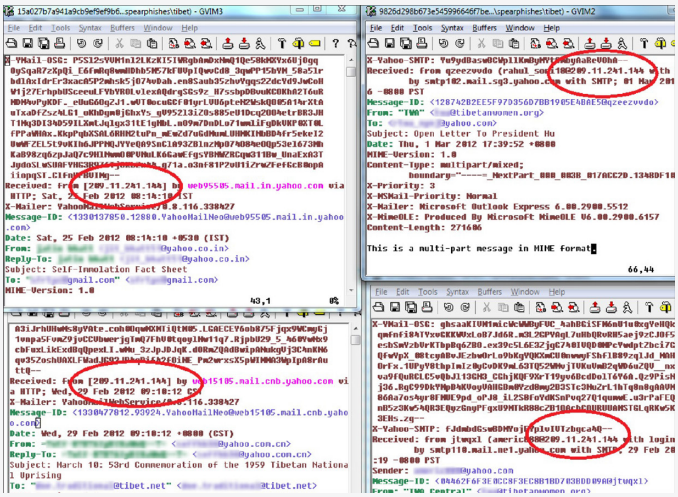
Alone, none of these attributes is absolute proof. But when multiple signs point to the same attacker, researchers can conclude with a high level of certainty who is behind a given campaign. That information can help anticipate attack methods and motivations, allowing security professionals to better anticipate future attacks and protect targeted systems and data.

Figure 15:
Four phishing emails



FireEye

Figure 16:
Phishing email heading information (IP addresses highlighted)



When faced with the urgent tasks of containing an attack and repairing the damage, determining the source of the attack might seem incidental. It is not. When a targeted organization knows the attacker's methods and objective, it can use that information to:

- Immediately shift resources to bolster vulnerable data
- Enlist additional help, whether internal resources or law enforcement
- More closely examine other vectors—possibly overlooked—that have been used by the attackers other campaigns

Knowing the source of an attack can be especially useful when combined with intelligence gleaned from previous attacks elsewhere from the same threat actor. Solutions such as the FireEye® Dynamic Threat Intelligence™ cloud—which shares anonymized threat intelligence across the growing base of FireEye customers—provides information about tactics, protocols, ports, and callback channels used by attackers.

To find out more about how the FireEye threat-protection platform can help you better defend against cyber attacks, visit FireEye at http://www.FireEye.com.

About FireEye

FireEye® has invented a purpose-built, virtual machine-based security platform that provides real-time threat protection to enterprises and governments worldwide against the next generation of cyber attacks. These highly sophisticated cyber attacks easily circumvent traditional signature-based defenses, such as next-generation firewalls, IPS, anti-virus, and gateways. The FireEye platform provides real-time, dynamic threat protection without the use of signatures to protect an organization across the primary threat vectors, including Web, email, and files and across the different stages of an attack life cycle. The core of the FireEye platform is a virtual execution engine, complemented by dynamic threat intelligence, to identify and block cyber attacks in real time. FireEye has over 1,000 customers across more than 40 countries, including over one-third of the Fortune 100.

© 2014 FireEye, Inc. All rights reserved. FireEye is a registered trademark of FireEye, Inc. All other brands, products, or service names are or may be trademarks or service marks of their respective owners. RPT.DB.EN-US.082014

FireEye