WALKING IN YOUR ENEMY'S SHADOW: WHEN FOURTH-PARTY COLLECTION BECOMES ATTRIBUTION HELL

Juan Andres Guerrero-Saade & Costin Raiu
Kaspersky Lab, USA & Romania

Email juan.guerrero@kaspersky.com;
craiu@kaspersky.ro

ABSTRACT

Attribution is complicated under the best of circumstances. Sparse attributory indicators and the possibility of overt manipulation have proven enough for many researchers to shy away from the attribution space. And yet, we haven't even discussed the worst-case scenarios. What happens to our research methods when threat actors start hacking each other? What happens when one threat actor leverages another's seemingly closed-source toolkit? Or better yet, what if they open-source an entire suite to generate so much noise that they'll never be heard?

Leaked documents have described how the standard practice of one espionage outfit infiltrating another has transcended into the realm of cyber in the form of fourth-party collection. While this represents an immediate failure for the victim intelligence service, the tragedy doesn't end there. Attackers can then go on to adopt the victim threat actor's toolkit and infrastructure, leveraging their data and access, and perpetrating attacks in their name. As interesting as this conversation could be in the abstract, we'd rather present examples from unpublished research that showcase how this is already happening in the wild.

Similarly, while we'd prefer to present threat intelligence research in its most polished and convincing form, fringe cases do appear. Strange activity overlaps between clusters, APT-on-APT operations, open-sourcing of proprietary tools, or repurposing of proprietary exploit implementations are some of the ways that the attribution and activity clustering structures start to break down and sometimes collapse. And this is not all an unintentional byproduct of our position as external observers; some threat actors are overtly adopting the TTPs of others and taking advantage of public reporting to blend their activities into the profiles researchers expect of other actors.

The material includes in-the-wild examples to substantiate previously hypothesized claims about attackers stealing each other's tools, repurposing exploits, and compromising the same infrastructure. These covert dynamics in the space of cyber espionage further substantiate the difficulties underlying accurate security research and the need to track threat actors continually. The examples we'll focus on come from unpublished research and unwritten observations from the original researchers themselves. The hope is to escape threat intel solipsism by providing a better framework to understand and discuss operations and actors and to understand how traditional espionage shadow games are being played out on the digital front.

INTRODUCTION

Opportunity plays a large and unaccredited role in the practice of intelligence. Where convenience can proffer information of intelligence gathering value, an intelligence service can capitalize while saving resources and maintaining a light touch to avoid detection. The maturity of an intelligence service can be measured in part by structural adaptation to maximize opportunistic collection wherever beneficial. Depending on the purview of the intelligence service in question, be it all source, human intelligence (HUMINT), signals intelligence (SIGINT), etc., these structural changes will take different forms. HUMINT outfits may recruit, supplant, or bug spouses, therapists, and priests in proximity to a desirable outfit to take advantage of the intended target's willingness to 'bare their soul' where they otherwise may not. The benefit of confession isn't strictly necessary; in some cases, the heat emanating from a facility, the number of pizza deliveries in a week, or the number of cars parked in a specific area may provide crucial information about the activities taking place in a location of interest. Creativity for opportunism is well rewarded in the practice of intelligence.

In the case of signals intelligence services, the gamut of opportunism extends wider still. By the very tradition of SIGINT collection, phone calls and emails are desirable, as are those of third-parties associated with the intended target who may not treat the content of those privileged interactions with the level of care of the original interlocutor. The existence of contacts and connections between targets of interest may be telling in and of itself. But let us widen our view to the options available to truly mature, sophisticated, and capable services when it comes to engaging the full berth of options at their disposal. After all, there's more than one SIGINT agency in the world. Different agencies, be they friend or foe, have a purview over different geographical areas and desirable sectors. Their analysts are likely to have a greater acquaintance with desirable targeting and the context in which to interpret the information received from their collection. This presents a valuable opportunity: to co-opt the collection methods of a foreign intelligence service to receive the same raw information being collected on targets of interest to the latter – or ideally both – intelligence services; this practice is known as *fourth-party collection*.

In the 21st century, intelligence operations of all kinds have embraced the ease, convenience, and tantamount desirability of digital espionage and as such, so must our concepts of intelligence methodology adapt. Fourth-party collection is an interesting and ongoing practice with a palpable impact on cyber espionage operations. The opportunities for fourth-party collection afforded by the digital realm are truly manifold. Sadly, so are the problematic implications for those of us interested in accurately researching or investigating these high-powered threat actors. From our perspective, the greatest difficulties lie in the realm of attribution where the first- and second-order implications of intelligence piggybacking threaten to destroy the current paradigm of our understanding and analysis capabilities. We must therefore carefully spell out our understanding of this practice, cases in the wild that showcase the *shadow of possible* fourth-party collection, and face the potential limits of our understanding of digital espionage to avoid missteps going forward.

A study of fourth-party collection from the perspective of outside observers is complicated. Passive forms of fourth-party

collection are largely unobservable from our vantage point. As such, we will discuss these insofar as we understand their theoretical effect on cyber operations. However, in the case of active fourth-party collection, with its heavy-handed byproducts like tool repurposing, victim stealing and sharing, we have endeavoured to provide in-the-wild examples alongside relevant thought experiments to best serve our understanding of this elusive practice.

Similarly, while we do not indict the SIGINT giants that protect our safety and the national security of our countries, we stand firmly in our understanding that techniques continue to trickle down and proliferate. While some fourth-party collection practices may be unattainable for actors that do not hold the status of 'gods on the wires', some of the complications that arise from code reuse or C&C (command-and-control) infrastructure piggybacking are already being felt and should be addressed with an objective view to furthering our understanding of the disastrous complications encountered therein.

Categorizing collection relative to its means of generation

The means by which information is generated and collected cannot be ignored when the purpose is to root out some semblance of actionable context that must not be tainted by the perspective of the source of the information or the limitations of the collection methods themselves. The analyst must be well aware of the means of generation and source of the information analysed in order to either compensate or exploit its provenance. For that reason, collection can be categorized by its means of generation in relation to the position of the parties involved, as discussed below. While insiders may find disagreement with the particulars of our appropriation of the following terms, these definitions will serve as functional categories for our understanding as outsiders looking into the more complex spheres of collection dynamics.

To facilitate this discussion we will fabricate a competent entity named 'Agency-A'[1] as a stand-in for a 'god on the wire'-style SIGINT agency interested in fourth-party collection. Let's categorize the collection categories available to this entity:

- ***First-party collection*** – Information collected by Agency-A first hand by passive or active means.
- ***Second-party collection*** – Information shared with Agency-A by partner intelligence agencies. Note that the original source of the data itself does not have to be the partner service's first-party collection.
- ***Third-party collection*** – Information collected by means of access to strategic organizations, be it wittingly, willingly, or otherwise. These may include Internet service and telecommunication providers, social media platforms, ad networks, or other companies that generate and collect vast amounts of data on potential targets of interest. Agency-A is in a position to correlate seemingly innocuous data (such as ad network trackers) with other forms of collection to benefit its overall intelligence product and targeting.
- ***Fourth-party collection*** – As described previously, fourth-party collection involves interception of a foreign intelligence service's 'computer network exploitation' (CNE) activity in a variety of possible configurations. Given the nature of Agency-A as a cyber-capable SIGINT entity, two modes of fourth-party collection are available to it: passive and active. The former will take advantage of its existing visibility into data in transit either between hop points in the adversary's infrastructure or perhaps in transit from the victim to the command-and-control servers themselves (whichever opportunity permits). On the other hand, active means involve the leveraging of diverse CNE capabilities to collect, replace, or disrupt the adversary's campaign. Both present challenges which we will explore in extensive detail further below.
- ***Fifth-party collection*** – This particular category is the 'unicorn of intelligence collection' and is more likely the result of serendipity rather than intentionally cunning access. For a fifth-party collection scenario to occur, Agency-A must be successfully conducting fourth-party collection on Agency-B, which, in turn, happens to be collecting Agency-C's respective collection efforts. Thereby, Agency-A's collection will entail a further level of opportunistic abstraction that yields the product of Agency-C's collection. While fourth-party collection is a desirable form of data collection, fifth-party collection is mostly a fortuitous byproduct that may instruct future tasking for Agency-A and likely not an intentional product in and of itself nor a reliable means of sustained production.

SIGINT SHOULDER-SURFING[2]

Each category of collection is interesting in its own right, but fourth-party collection in particular is worthy of added interest and scrutiny from the threat intelligence research community. This heightened status is warranted by the troubling ramifications implicit in the exploitation of 'cyber situational awareness' by any entity involved in fourth-party collection. Unwarranted tin-foil-hat-worthy claims of prevalent false-flagging become marginally[3] more plausible when modelling for threat actors in the rare position to conduct fourth-party collection. We will first discuss the situational models that enable fourth-party collection, followed by the byproducts and benefits of this collection category, before discussing the specific problems these pose for threat intelligence research.

Passive collection

The importance of 'god on the wire'[4] status for our hypothetical Agency-A becomes most immediate when discussing passive collection opportunities. These are

[1] Similarly, we will use Agency-B as a second, semi-competent SIGINT agency upon which Agency-A can be recurringly predatory for the sake of explanation. When necessary, an even less competent entity, Agency-C, will serve as prey.

[2] The following section will build on the elementary model of fourth-party collection described in the leaked slidedeck 'Fourth Party Opportunities – I drink your milkshake' [1].

[3] Keeping in mind that many of these threat actors are more likely impaired by a litany of lawyers than propelled forward by audacious tasking and attributory cyber-acrobatics.

[4] By 'god on the wire' status, we are referring to the sort of entity that has regular and legitimate access to nationwide, international, or even transcontinental taps providing them with otherwise inimitable access to data in transit over a given region.

characterized by a silent ability to collect data at some point in transit between the victim and the original attacker (Agency-B).

For those unfamiliar with the fundamental elements of a cyber espionage campaign, most consist of a malware toolkit meant to perform victim tracking and collection of specific data on a periodic and persistent basis. Like a satellite beaming back to Earth, the malware exfiltrates this information back to a command-and-control server. This communication usually involves the retrieval of further tasking or instructions from the attackers, or perhaps next-stage malware. As such, we can expect the attackers to connect to this command-and-control server (or a daisy-chain arrangement of servers) to retrieve exfiltrated information, upload tasking and payloads, and generally monitor their infrastructure.

This snapshot of a standard cyber espionage operation can function as a working model for how we can expect Agency-B to conduct its day-to-day operations. Where, then, does Agency-A come in? Depending on its existing access, Agency-A can leverage multiple passive collection opportunities, as follows:

Victim-to-server transit

By virtue of its placement, Agency-A may have access to the data in transit from the victim to one or more of Agency-B's command-and-control servers. Possible enablers include: taps on the cables leading to these servers, access to routing infrastructure along the route, or broad third-party access to the Internet Service Providers (ISPs) along the way or the Virtual Private Server (VPS) hosting provider specific to this server.

Server-to-attacker transit

Alternatively, Agency-A can find a similar vantage point along the route connecting Agency-B and its command-and-control (C&C) infrastructure, or within junctures connecting different nodes in that infrastructure itself. The latter is possible within C&C configurations that employ server daisy-chaining for 'anonymizing' properties. In that case, multiple servers are set up as relay hops, employing encryption and a fragmented awareness of the full route of servers as means of protecting the operators. For example, the implant on the victim machine will exfiltrate data to Server-P. Server-P is aware of Server-Q (and only Server-Q) and will proceed to forward the victim's data there. Server-Q will in turn forward to Server-R, and so on. By placing these servers across many jurisdictions and different hosting providers, attackers with middling sophistication hope to throw off law-enforcement and researcher attempts to track or disrupt their activities. However, this system of relays actually provides multiple potential opportunities for Agency-A to passively collect this traffic in transit from one node to another, dependent on Agency-A's access and luck.

Decryption and deobfuscation

Assuming that Agency-A possesses the capability to decrypt or deobfuscate Agency-B's means of data protection in transit, it will be capable of reaping the benefits of Agency-B's targeting passively and without arousing suspicion. For those ardent fans of encryption that may consider this unlikely, it's worth emphasizing that obfuscation methods chosen by malware developers of this calibre are far from infallible. In cases where potent encryption has been adopted, this encryption is only as strong as the means of storage of the decryption keys in what is likely an automated system somewhere in Agency-B, or worse yet somewhere in Agency-B's public-facing attack infrastructure. Wherever costly superhuman cryptanalysis capabilities appear necessary, access to a specific endpoint[5] is ever more likely to play a cost-effective and desirable means of success.

Compromised servers, dynamic providers, open-ownership servers, and abused public services

As may appear obvious by now, the misuse of public resources or compromised servers is likely to yield Agency-A several fourth-party collection opportunities as well. Where we have thus far characterized Agency-B as a mid-range attacker with some sophistication due to its deployment of reasonable countermeasures and targeting ambitions, we can point to Agency-C as an example of the now all too common low-resource threat actor entering the digital espionage space. Agency-C-style attackers often employ trifle means, script-kiddie-worthy tools, and truly unimaginative infection vectors. However, lack of sophistication has not kept these attackers from succeeding in their collection efforts.

Agency-C is likely to attempt to 'hide in the noise' while keeping costs low, particularly when it comes to its infrastructure. By compromising vulnerable servers or abusing public resources like cloud providers, open proxies, and dynamic DNS services, Agency-C may hope to mix in with an abundance of promiscuous attackers more likely to be ignored by high-calibre enforcers. However, in reality, Agency-C is likely to provide Agency-A with fourth-party collection opportunities galore by virtue of well-placed third-party access to the more reputable hosting services. These service providers are even less likely to balk at the use of this access considering that the target is abusing the services of a well-guarded corporate brand to conduct malicious activities.

Active collection

So far we have focused entirely on passive collection capabilities that are likely to yield a benefit for Agency-A while providing little to no indication of its access. Active collection requires a heavier hand but may be the only recourse when convenient opportunities for passive collection have not presented themselves. An alternative compelling rationale states that active collection is likely to yield greater returns for the overly-capable opportunistic attacker interested in greater cyber situational awareness.

Active collection involves CNE on systems that form part of the target campaign's command-and-control infrastructure. The idea is for Agency-A to break into one or more of Agency-B's command-and-control servers, or better yet a backend-collection node. Though access with stolen credentials may suffice on a temporary basis, a disciplined attacker is more likely to opt for a more regimented and reliable form of access. An attacker might consider backdooring a key program on the server, like SSH, thus enabling persistent access without arousing suspicion. The

[5] By means of Agency-A's well-targeted CNE.

advanced attacker who realizes the collection value of the server is more likely to place their own stealthy implant on the C&C to collect and exfiltrate information collected from victims regularly as well as monitor the original operators connecting to these servers.

An equation for heavy-handed opportunism

The idea of monitoring the original operators as they connect to these nodes is crucial considering that a key starting point for fourth-party collection is exceptional cyber situational awareness[6]. The apex predator looking to exploit the cyber operations of lesser equipped threat actors must be aware of their existence, their procedures and techniques, their tasking, the effectiveness of their countermeasures, and so on. This detailed situational understanding is then coupled with a measure of what may be termed Reasonable Attainable-Access Potential (RAAP) and weighed in relation to what stands to be gained from this style of collection.

We introduce the notion of RAAP with a particular interest in understanding the logistics of a 'god on the wire'-style SIGINT agency with a cyber remit. Though these are unequivocally placed at the very top of the adversary pyramid, they are not, in fact, omnipotent nor do they possess unlimited resources, and as such must be understood with an eye towards logistical realism. RAAP speaks to the measure that determines whether the disposition to gain access to a target can materialize within reasonable means.

Barring the paramount value of a truly critical target, chances are that even an apex predator with vast resources will not take the time to tailor brand new, time-consuming, and costly exploits and implants for a one-off victim (or an unusual platform) without a drastic return on investment. Therefore, we can expect the path of least resistance to be determined along the following line of questioning:

- 'Are any of these systems vulnerable to exploits we already possess?'
- 'Do we already have an implant suitable for this platform?'
- 'Are credentials already available from other collection sources?'
- 'Might we have access to a relevant critical juncture by means of a third-party?'
- 'Does an allied service (second-party collection) already possess access to our intended victim?'

We must be prepared to consider that even a desirable (but not critical) target may fall off the tasking list entirely if these considerations prove even marginally unfavourable. The tragic banality of public sector decision-making entails that a simple lack of serendipitous bureaucratic support is likely to play a heavy hand in prioritizing one opportunity over another.

A server for two masters

What we should expect to see in practice are command-and-control servers that are either backdoored to provide a persistent means of access, or an autonomous implant (whose ownership is unrelated to the operation's stakeholding threat actor) collecting information from that box. To better clarify the situation, let's apply this concept specifically to dedicated VPSs operating as C&Cs and not compromised servers where ownership ambiguity implicitly applies.

Managing a sprawling C&C infrastructure is a costly and difficult problem, resistant to perfect automation, prone to failures, misconfigurations and lapsed registrations. C&C management is the cornerstone of the thesis that cyber operations do not scale well – hence the proliferation of operational security failures and mistakes[7]. As threat actors become favoured by their tasking entity (i.e. sponsor government), their operational requirements are often drastically increased without expanded resources or adequate time for tooling and preparation.

Despite their involvement in compromising systems, campaign operators are often as unconcerned about the integrity of their own systems as their victims. Talent in these areas is non-transferable: those devious operators endowed with the gift of offensive capabilities are no closer to solving the issue of maintaining device integrity than we are of perfecting our offensive capabilities by virtue of engaging in security research. To that point, despite being managed by proficient attackers, the inner workings and network interactions of command-and-control servers are in many cases no better monitored than any other system.

Let us envision the following idealized hypothetical situation:

Agency-A becomes aware of Agency-B's sprawling campaign in desirable territory Q. As Agency-A is not aware of the particular targets that would prove desirable within that territory, it proceeds to map Agency-B's campaign infrastructure, noting that none of the nodes are particularly accessible by means of passive collection alone. Assessing the Reasonable Attainable-Access Potential of the servers involved, it turns out that Agency-B's servers are reasonably accessible: either by means of a vulnerability for which Agency-A has already developed an exploit (or could abuse with little to no investment), or perhaps by means of credentials already[8] within Agency-A's databases. But successful initial access represents an ephemeral beachhead for our idealized attacker.

Having established initial access, a gung ho actor might set off to explore the server's pilfered contents like a kid in a candy store, but the prolific Agency-A will instead seek to understand the intended set-up of the server: *does it function as a relay for the operators – or as a staging server for tools and payloads? Does it host exfiltrated material, or does it forward it to a chain of other servers?* Barring the rare scenario where the operation of that server might involve some stringent security practices, Agency-A will seek to maximize the value of its access relative to the determined

[6] The term 'cyber situational awareness' refers to an institution's awareness of the cyber operations of other actors in a relevant region (or ideally at a global scale). While this is a desirable attribute for corporations and public-facing value stakeholder entities, it is a crucial ingredient for fourth-party collection as it constitutes the fundamental insight leveraged to target desirable nodes and stay ahead of the victim intelligence service's methods.

[7] The sort of fortunate mistakes on which threat intel researchers can capitalize to hunt for further elements in a campaign, or perhaps even gain greater information on the identity of the actor involved.

[8] Agency-A may already possess relevant credentials by means of automated bulk passive collection systems, second-party collection or all-source intelligence sharing, third-party collection with access to relevant service providers (for example, the email provider used during a VPS registration, or the VPS hosting provider), or simply by means of previous acquaintance with Agency-B's poor infrastructure management practices (after all, password reuse cuts both ways).

functionality of the server. At the very least, Agency-A will likely place a webshell or replace a system binary with one that ensures persistent, responsible[9] access going forward.

If the functionality of the server signals high projected value to its strategic interests, Agency-A is likely to upgrade this access with an implant that automates the desired fourth-party collection. This implant's functionality can be as simple as forwarding a copy of all exfiltrated materials to Agency-A's own collection infrastructure. In all likelihood, the opportunity will be maximized to automate greater capabilities still, including:

- *Logging of all operator activities*, whether in the case of a relay server or to understand operator tasking and exfiltration practices.
- *Server-side beaconing*, particularly useful in the case of infrastructure that's continually reassigned or taken on- and off-line and only made available during attack phases.
- *Watermarking of files, placing callback beacons or further payloads* within pilfered materials to gain tailored access to the operator's machines.
- *Tool collection*, particularly useful in the case of a staging server that will contain malware and exploits intended for use in current and future campaigns.

Furthermore, in order to protect Agency-A's access, the implant may be loaded with a rudimentary AV-like component. The intended functionality is multi-dimensional: by monitoring for the presence of other known – or as of yet unknown – adversaries, Agency-A protects its access, the integrity of its investment (by protecting a likely unknown implant from being discovered and analysed), and increases its cyber situational awareness capabilities by learning more about other threat actors that have reached a similar apex stage of maturity to conduct similar operational manoeuvres[10]. The functionality consists of monitoring for known adversaries' indicators of compromise and the generation of general or specific telemetry [2] from the victim server. After all, it's the fourth-party collector's curse to look over its shoulder as well, lest it fall prey to its own tactics.

'We heard you like popping boxes, so we popped your box so we can watch while you watch'

As was clearly stated at the outset of this research endeavour, attempting to highlight examples of fourth-party collection is a difficult exercise in the interpretation of shadows and vague remnants. While passive collection is beyond our ability to observe, active collection involves the risk of leaving a footprint in the form of artifacts. In the course of APT investigations, *Kaspersky Lab*'s Global Research and Analysis Team (GReAT) has encountered strange artifacts that defy immediate understanding in the context of the investigation itself. While we cannot be certain of the intent or provenance of these artifacts, they nonetheless fit a conceptual framework of active fourth-party collection and are presented as such:

Crouching Yeti's pixelated servers

In July 2014, we published our research [3] on Crouching Yeti, also known as 'Energetic Bear', an APT actor active since at least 2010. Between 2010 and 2014, Crouching Yeti was involved in intrusions against a variety of sectors, including:

- Industrial/machinery
- Manufacturing
- Pharmaceutical
- Construction
- Education
- Information technology

Most of the victims we identified fell into the industrial and machine manufacturing sector, indicating vertical of special interest for this attacker.

To manage its victims, Crouching Yeti relied on a network of hacked websites which acted as command-and-control servers. For this, the attackers would install a PHP-based backend that could be used to collect data from or deliver commands to the victims. To manage the backend, the attackers used a control panel (also written in PHP) that, upon checking login credentials, would allow them to manage the information stolen from the victims.

In March 2014, while investigating one of the hacked sites used by Energetic Bear, we observed that for a brief period of time, the page for the control panel was modified to include an <img src> tag that pointed to a remote IP address in China. This remote 1x1 pixels wide image was likely intended to fingerprint the attackers as they logged into their control panel. The fingerprinting could have been used to collect attributory indicators. The usage of an IP address in China, which appeared to point to yet another hacked server, was most likely an attempt at a rudimentary false flag should this injection be discovered.

NetTraveler's most leet backdoor

While investigating the NetTraveler attacks, we obtained a disk image of a *mothership server* used by the threat actor. The mothership, a combination staging and relay server, contained a large number of scripts used by the attackers to interact with their malware, as well as VPN software and other IP masking solutions used to tunnel into their own hacking infrastructure.

Beyond the fortuitous boon of seizing such a content-rich server, GReAT researchers made a further unexpected discovery: the presence of a backdoor apparently placed by another entity.

We believe the backdoor was installed by an entity intent on maintaining prolonged access to the NetTraveler infrastructure or their stolen data. Considering that the NetTraveler operators had direct access to their mothership

[9] The idea of responsible access was sparsely debated (and unfairly criticized) by the information security community under the term of art 'NOBUS' (No One But US) backdoors.

[10] Nothing discounts the possibility that an implant in a vulnerable command-and-control server may have been placed there by the newer player in the fourth-party collection space: the more reckless security researchers. While most companies will not condone breaking into command-and-control servers (much less interfering with the integrity of servers that are possible subjects of ongoing law-enforcement investigations), it's nonetheless not beyond the pale for an overzealous threat intelligence company to enable some form of permanent access to continue to monitor a threat actor's operations.

server and didn't need a backdoor to operate it, we consider other possible interpretations less likely.

The artifact encountered is the following:

Name	svchost.exe
MD5	58a4d93d386736cb9843a267c7c3c10b
Size	37,888

Interestingly, the backdoor is written in assembly language and was injected into an empty Visual C executable that served as a template. This unusual implementation was likely chosen in order to confuse analysis or prevent detection by simple anti-virus programs.

The backdoor is primitive and does nothing but listen to port 31337[11] and wait for a payload to be sent. The acceptable payload format is depicted in Figure 1.

Figure 1: Acceptable payload format.

The assembly code is then executed and can perform any action chosen by the predatory attackers. The backdoor requires no authentication. Combining this sort of backdoor with Metasploit or other similar frameworks could easily have been used to control the system.

Black sheep wall

In June 2016, *Kaspersky Lab* researchers discovered an unknown zero-day *Adobe Flash Player* exploit actively leveraged in targeted attacks. Further analysis revealed payload overlaps with the DarkHotel threat actor. DarkHotel is known to have deployed several *Adobe Flash Player* exploits over the years. What makes this case particularly interesting is the fact that one of the websites compromised by DarkHotel for use in watering hole attacks hosted exploitation scripts from another APT group. We code-named this second actor 'ScarCruft'.

According to our telemetry, ScarCruft appears to have been targeting Russian, Chinese, and Korean-speaking companies and individuals, among others. This actor relies on watering hole and spear-phishing attacks to infect its victims.

The most interesting overlap between DarkHotel and ScarCruft became apparent with two operations we named 'Operation Daybreak' and 'Operation Erebus'.

Operation Daybreak appears to have been launched by ScarCruft in March 2016 and employed a previously unknown (zero-day) *Adobe Flash Player* exploit. The script used for exploitation was hosted at the following link:

```
hxxp://scarcroft[.]net/plus/thumbs/index.php
```

At the end of May 2016, *Kaspersky*'s advanced heuristic detection technology caught a new, unique web attack abusing the CVE-2016-4117 vulnerability. The malicious payloads were distributed from compromised websites and didn't display apparent connections to previously known malware. We decided to call this 'Operation Erebus'.

In Operation Erebus, the two hacked websites used in the attacks included the following link:

```
hxxp://scarcroft[.]net/wp-content/plugins/twitplug/
twitter.php
```

Additional links included:

```
hxxp://www[.]chateau-eu[.]fr/wp-content/player/
qqplayer.php?...

hxxp://www[.]chateau-eu[.]fr/wp-content/player/
qqplayer.jpg

hxxp://www[.]chateau-eu[.]fr/wp-content/plugins/
gallery/photo-gallery.php?...

hxxp://www[.]chateau-eu[.]fr/wp-content/protect/wp-
protect.php?...
```

Those links delivered a CVE-2016-4117 exploit, ripped from previously known samples delivering FinFisher payloads, with a slightly modified shellcode and payload URL, as shown in Figure 2.

[11] The most 'LEET!' port.

Figure 2: The links delivered a CVE-2016-4117 exploit, ripped from previously known samples delivering FinFisher payloads, with a slightly modified shellcode and payload URL.

The exploits were delivered through watering hole attacks from several compromised websites, including:

- `rfchosun[.]org`
- `dailynk[.]com`
- `cafe.daum[.]net`

Figure 3 is a diagrammatic representation of the attacks, which serves to better illustrate the connections and overlaps.

According to our telemetry, DarkHotel's Operation Daybreak links were used as early as April 2016; ScarCruft's Operation Erebus links were first used in attacks on 26 May 2016. This suggests the possibility that the ScarCruft actor may have observed the DarkHotel attacks. They succeeded in breaching the same website and used it for another set of attacks on 26 May. The hacked site overlap was enough to trick us (and other researchers) into believing that ScarCruft and DarkHotel were the same threat actor, and thereby that Operation Erebus and Operation Daybreak were launched by the same actor. It is now clear that this is not the case.

Focusing on TTPs and victim tasking is useful in further disambiguating the two threat actors (Figure 4).

DarkHotel's Operation Daybreak (Figure 4, right) relied on spear-phishing emails served to victims in the geographical focus illustrated in Figure 4: predominantly targeting Chinese victims with a *Flash Player* zero-day. Meanwhile, ScarCruft's Operation Erebus focused primarily on South Korea.

More recently (as of April 2017), the ScarCruft actor has been attacking South Korean institutions for both espionage and sabotage in relation to the presidential elections. While the main goal of this recent wave of attacks is stealing information from valuable targets, it appears that creating social chaos is not beyond the ScarCruft actor's intent from time to time. For this purpose, ScarCruft leveraged malicious *Hangul* documents (.HWP) that drop a destructive payload. Unlike the ScarCruft actor, DarkHotel has not been observed engaging in destructive operations to date.

BYPRODUCTS OF NON-CONSENSUAL INTELLIGENCE SHARING

Having detailed the means and constraints that enable and guide the practice of fourth-party collection, we'd do well to understand its benefits and byproducts. This allows us to take into account the incentives that justify this intricate, delicate,

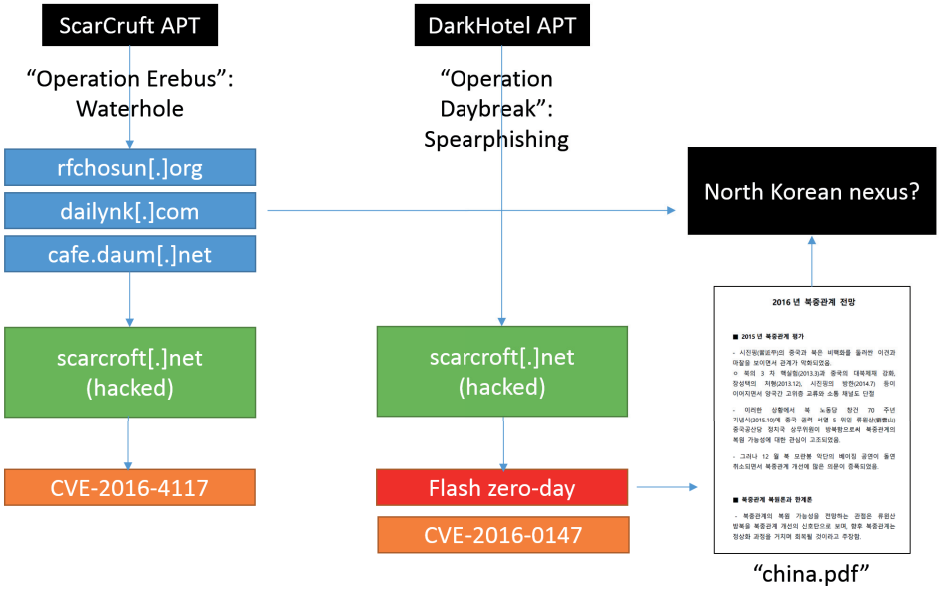

Figure 3: Diagrammatic representation of the attacks.

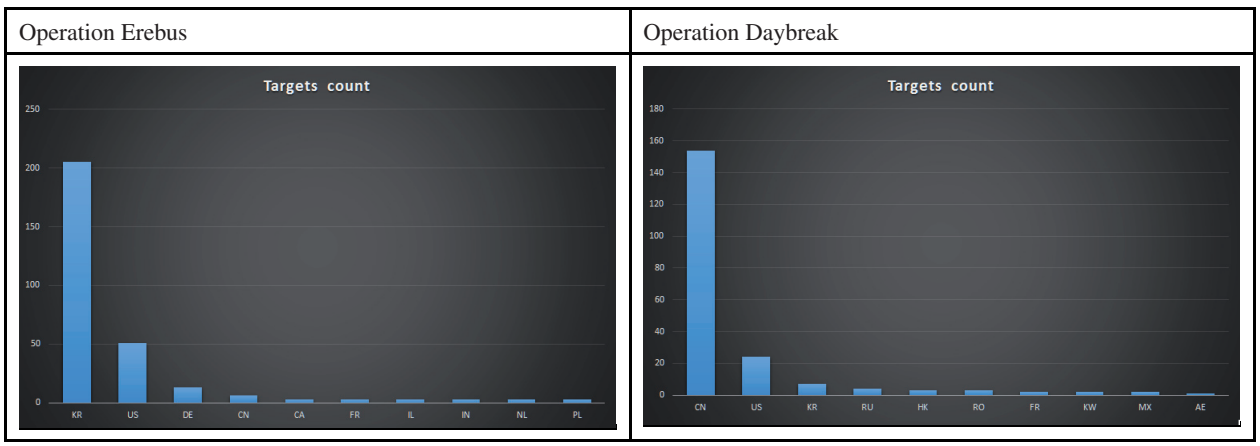

Figure 4: Focusing on TTPs and victim tasking.

and investment-heavy practice. The main aspects we'll focus on are tasking, code reuse, and adversarial learning benefits. Where possible we have added small but illustrative examples encountered in the wild.

Tasking-by-proxy

Tasking is one of the most revealing aspects of an operation. Quality tasking (i.e. avoiding a 'spray-and-pray' approach) is the mark of a mature and measured threat actor. Few reach this level of care, as most continue to rely on wide-ranging spear-phishing waves followed by lateral movement intended to map the network-to-individual topology of a breached organization. Some services may opt for costly HUMINT or inexact OSINT measures. However, for the most capable SIGINT agencies, fourth-party collection represents a way to avoid this costly and noisy phase altogether. This is accomplished by taking tasking queues from other threat actors – particularly those with a stakeholder role over a desirable region or organization: the more capable Agency-A is able to pinpoint valuable systems and individuals by virtue of Agency-B's prolonged interest in them and the observable quality of the exfiltration or functional value evident in Agency-B's resulting collection.

Not only is Agency-A able to lower its investment threshold for its own campaign in a foreign region thanks to fourth-party collection, it may also be able to leverage another threat actor's access to further its own access. The more intimate Agency-A's understanding of the techniques and tools leveraged by a lesser-grade attacker, the more likely opportunities for victim stealing will present themselves. For example, if Agency-B is using an implant that accepts next-stage payloads without stringent checks for provenance, Agency-A can place its own implant by means of the careless design of Agency-B's implant. Or perhaps Agency-C is so oblivious as to allow its implants to accept commands promiscuously, allowing Agency-A direct operator access to accomplish further access. These, and other mistakes, would allow Agency-A to swoop in and steal a foreign service's victims altogether.

Agency-A can not only place its own implant in the victim box but also proceed to clean out the foreign malware. We have witnessed interesting cases where advanced implants were leveraged as benevolent protection solutions for likely oblivious high-value religious figures under constant attack by hostile governments.

Alternatively, threat actors may well choose to share a victim. This may well happen if another threat actor is an ally or if Agency-A understands that the other threat actor inhabiting the same box is so persistent as to likely insist on maintaining or regaining access. By allowing another to exist on the same victim box, Agency-A maintains its access without arousing suspicion from another capable service as to the presence (and characteristics) of Agency-A's toolkit. However, it must be noted that this is high-risk behaviour.

Where an actor of the quality and diligence of Agency-A is unlikely to arouse suspicion by design, and thereby won't provide many opportunities to get caught, reckless lesser-grade actors will likely arouse eventual scrutiny. When an incident response team shows up to weed out the loud and careless Agency-C, those same forensic artifacts are likely to contain traces of Agency-A and any other intelligence service that may have been inhabiting the same high-value target. If this seems unlikely, keep in mind that tasking does not occur in a perfectly isolated theoretical vacuum. Tyrants, political figures, and high-value researchers use one or multiple systems, and a collection agency tasked with gaining as much information on these figures as possible has a limited number of opportunities to do so consistently. The avid reader may remember the story of the 'Magnet of Threats': a machine in the Middle East of such high value as to simultaneously host implants for Regin, Turla, ItaDuke, Animal Farm, Careto, and Equation. It was an investigation into this very system that yielded the first sighting and discovery of the prolific Equation Group.

Why reinvent the wheel?

While we may subdivide artifacts into exploits, infection vectors, implants, etc., they're ultimately tools for operators to accomplish specific goals as they carry out their campaigns. As such, who's to blame operators for stealing each other's tools? One of the possible boons of fourth-party collection is access to staging servers that may host exploits, implants, scripts, and possibly even source code (depending on the carelessness of the victim service). Exploits, particularly zero-days, are of immediate benefit for any threat actor and are likely to be ceased and repurposed immediately to maximize the ongoing pre-patch window. Tools and implants pose a more insidious gain and one that has caused great (if perhaps overblown) concern in the infosec community's sadomasochistic flirting with armchair attribution.

The more recent point of contention is that of code reuse and repurposing. For experienced malware analysts, clustering[12] malware on the basis of code reuse is a golden practice, with both industry standard and proprietary tools that can match function overlaps down to a percentage point. Partial code similarities or reuse are not the beginning of an unsolvable paradox as popular Tweets might have us believe. Instead, they suggest that the developers chose to implement a portion of code to which they had access (whether originally theirs, from a forum, a book, or wherever) once again. This form of clustering continues to stand up to scrutiny and presents great value for campaign and threat actor understanding, within the bounds of reasonable insight that clustering – *and not attribution* – can provide.

Recent PR-worthy global incidents were impacted by insights gained largely thanks to code reuse that allowed for threat actor clustering. The two incidents that come to mind are the WannaCry cryptoworm and SWIFT hack clustering to the Lazarus Group and their BlueNoroff subset, respectively.

The original research[13] that yielded the bulk of the Lazarus Group cluster was made possible entirely by the threat actor's consistently careless cut-and-splice methodology for managing a gigantic codebase in order to yield different tools

[12] Clustering is an important term in malware analysis that refers to the ability to group similar artifacts together. It's a term the wider infosec community would do well to avail itself of, rather than obsess over attribution, which looks to cross the fifth domain's boundary to pinpoint a perpetrating organization or group of individuals.

[13] Joint research between Juan Andres Guerrero-Saade (*Kaspersky* GReAT) and Jaime Blasco (*AlienVault*). Presented in parallel with *Novetta*'s tour de force in profiling nearly a decade's worth of Lazarus Group tools under the moniker 'Operation Blockbuster'.

and malware families as needed. Despite this ideal ground for clustering, many were left unsatisfied by the findings as they were recast as clustering-cum-attribution claims.

Once we move away from the temptation to equate clustering and attribution, we can clearly identify the issue of code repurposing of privileged, proprietary, and closed-source code as one that challenges our ability to accurately and consistently cluster threat actors. If Agency-A has access to Agency-B's source code and begins to splice the latter's functionality into its own malware, how do we manage the clustering overlap? What about when Agency-A decides to run an operation entirely using Agency-B's source code (but perhaps relying on an entirely different infrastructure)? What if Agency-C were to get a well-placed human asset in Agency-A to quietly steal the far more prolific actor's tools and begin to leverage them as its own? Malware clustering yields straightforward results but threat actor clustering (and subsequent attempts at interpretative attribution) fray and suffer in the face of fourth-party collection gains as they always have in the face of HUMINT infiltrations and operator defections.

A minute of schadenfreude, a year of tooling

A perhaps less controversial example of code repurposing can be observed in the aftermath of the poorly disseminated *HackingTeam* dump. While the latter commanded great public interest into the company's contracting practices with objectionable governments, the dumping of already weaponized zero-day exploits and a full malware codebase posed greater gains for determined attackers than the benefits it conferred on Internet safety. For example, DarkHotel was seen repurposing a *HackingTeam* Flash exploit [4] mere days after the dump. And while the malware codebase itself may prove too intricate a set-up to mount in its entirety, determined attackers have found ways to make the implant code useful in the wild.

For those unfamiliar with the infamous developers, *HackingTeam* is a Milan-based information technology company that sells offensive intrusion and surveillance capabilities primarily to governments and law enforcement agencies. Its main product is called *Remote Control Systems* (*RCS*) and enables monitoring of the communications of Internet users, deciphering their encrypted files and emails, recording *Skype* and other VoIP communications, and remotely activating microphones and cameras on target devices.

HackingTeam's *RCS* features a multi-stage attack platform that relies on several different trojans in order to minimize the chances of discovery and operational loss. The first-level validator-style trojans are referred to as 'Scouts'. If the victim is confirmed as the designated target, the 'Scout' is upgraded to one of the more sophisticated implants, either 'Soldier' or 'Elite', with each level adding further features and capabilities to spy on the victim.

'Dancing Salome' is the *Kaspersky* codename for a previously undisclosed APT actor with a primary focus on ministries of foreign affairs, think tanks, and Ukraine. What makes Dancing Salome interesting and relevant is the attacker's penchant for leveraging *HackingTeam RCS* implants compiled after the public breach. The name 'Dancing Salome' was derived from a hard-coded directory in the *RCS* binaries, <X:\RCS\salome\WinWord.exe>. A subset of Dancing Salome-related activities was subsequently reported as 'Operation Armageddon'.

In 2016, we observed two previously unknown *HackingTeam RCS* ('Scout') implants uploaded to a multi-scanner service:

Name	Upload date	Country of upload	Hash
Security Conference Agenda.docx	12 Feb 2016	Russia	99da44b0f1792460ba1d6730c65ec190
Invitation.docx	27 Jan 2016	Poland	d1f1d7b84bb5bc84243c3b43e93622cd

Table 1: Previously unknown HackingTeam RCS implants.

The implants were embedded into *Microsoft Word* files. The first document refers to a conference about 'Domestic Developments in the South Caucasus'. The second document is an invitation to an event on 'South Caucasus Security', allegedly organized by the ConcorD Centre for Political and Legal Studies in cooperation with Friedrich Ebert Stiftung [5], a non-profit German political foundation committed to the values of social democracy. According to its website:

> *'The Friedrich-Ebert-Stiftung (FES) is the oldest political foundation in Germany, with a rich tradition in social democracy that dates back to its founding in 1925. The foundation owes its formation and mission to the political legacy of namesake Friedrich Ebert, the first democratically-elected President in German history.'*
>
> *'The Regional Office South Caucasus is based in Tbilisi and coordinates programs in Georgia, Armenia and Azerbaijan. In addition, there is a Liaison Office in Yerevan that focuses on projects in Armenia under the umbrella of the Regional Office in Tbilisi.'*

These documents do not contain exploits; instead, they rely on the user clicking on the embedded OLE objects inside.

The OLE storage in the *Word* documents holds the final *Windows* PE payloads, which have the following identifying data:

- `6355c82c7c6a90ef41824a03bbabbabc`
- `99a18bf3c04a491b256f7d60eb6e0f26`

Both executables are VMProtect-ed *RCS* 'Scout' binaries. The two Scout binaries used in the Dancing Salome attacks are signed with an invalid digital certificate issued to 'SPC'[14]. Both samples have been configured to work with the same command-and-control server: 89.46.102[.]43.

An interesting indication that Dancing Salome is not a legitimate *HackingTeam* customer comes from the samples' hard-coded customer ID. As a means of tracking samples that might leak to multi-scanners, all *HackingTeam* samples contain a customer ID that points to the licence of the *RCS* MasterNode used to generate them.

In the case of Dancing Salome, the customer ID can easily be spotted in the decrypted payload (Figure 5).

[14] Serial: 58 be 7b 63 89 43 cb 90 44 1a 09 42 25 ed c5 1c.

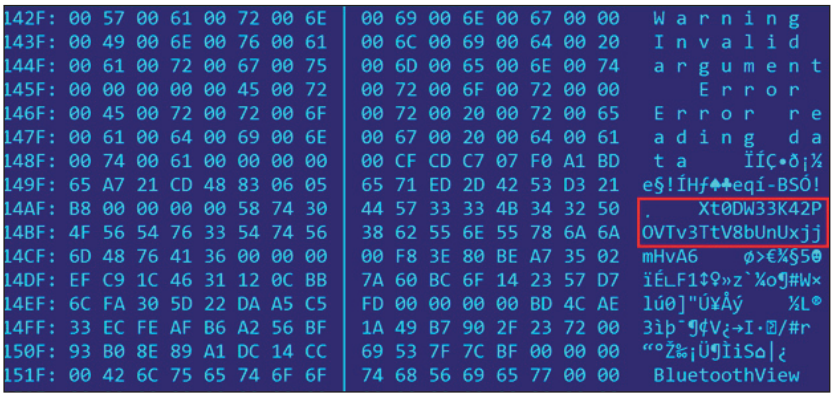

Figure 5: HackingTeam customer ID string inside decrypted samples.

In this case, the customer ID is 'Xt0DW33K...'. A dedicated tool called rcs-kill.rb [6] released with the code leak allows us to identify the customer (or licence) name based on the string shown in Figure 6.

```
"GDWwVyrq"=>"dod",
"Kwh80g9E"=>"edq",
"Xt1DW33K"=>"fae-demo",
"Xt0DW33K"=>"fae-poc",
"j84fj1Ej"=>"gedp",
"vIByzgbS"=>"gip",
"KY4pBxoC"=>"gnse",
"nFGPKB8T"=>"ida-prod",
"HtAUfHdq"=>"ida-test",
```

Figure 6: We can identify the customer or licence name based on this string.

In this case, 'fae-poc' means 'field application engineer – proof of concept'. Field application engineers are *HackingTeam* terminology for those in charge of on-site customer demos of *RCS*. In the console-to-mainframe design of *RCS*, field engineers can connect to a MasterNode (malware factory) server with limited licensing in order to generate backdoors that can be tested for clients. The use of this particular designation may imply that the attacker has stood up a minimalist MasterNode installation and is using the proof-of-concept mode to generate these binaries. This may happen either through the distribution of unlicensed software or, more likely, as re-compiled source code from the *HackingTeam* dump itself.

To further support this theory, some samples contain strings referencing a BleachBit executable used as one of the default [7] *HackingTeam* Scouts for field engineers to showcase to customers. Of course, purchasing a fully featured *HackingTeam RCS* system is quite expensive; in this case, we believe the attackers repurposed the leaked *RCS* sources in a truly cost-effective manner. While the clustering confusion is ultimately mitigated due to the actor's inability to wield the *HackingTeam* malware suite effectively and their reliance on new infrastructure, it stands as a clear example of how access to a privileged codebase can problematize threat actor clustering based on code reuse alone.

Offensive leveraging of cyber situational awareness

It's only natural that an actor that has found a way to effectively weaponize their existing cyber situational awareness for great gains will know to value the generation of further awareness of the activities of other capable threat actors. There are many methods by which to hunt and track new threats in cyberspace. Different research outfits have developed their own styles to match the peculiarities of the data sources available to them. SIGINT giants with an information assurance or government-wide cyber defence remit have also developed impressive capabilities to leverage their unique visibility to track adversaries. External observers often consider cybersecurity a commodity state accomplished by the acquisition of expensive combinations of hardware and software. True practitioners actually plagued by the ghastly task of defending a sprawling heterogeneous network have instead accepted that, not only will intrusions happen and incident responders be needed, but that so many of these incidents will take place so often that an awareness of existing threats must guide how they leverage the very limited resources available for this task.

The nascent private offering of threat intelligence is designed to provide a context to guide this complex calculus. It says, *'you have adversaries, these are known to other organizations facing similar threats, here are their past and current tactics, tools, and infrastructure, defend and hunt accordingly'*. Most threat intelligence producers are servicing a variety of customers in different verticals, with different capital and value features, and so on. As such, producers are not in a position to genuinely identify what threat actors will be relevant to their customers (or the public interest as a whole, in the case of those altruism/PR-driven producers still publishing indiscriminately). This is the main thrust behind our continual argument that the only genuine threat intelligence is global in scope[15].

While mere mortal threat intelligence producers are barred from greater prescience regarding the ultimate applied

[15] A thought-experiment: A threat intelligence producer contracted by Agency-A encounters Agency-A's toolkit during a separate incident response engagement in an unrelated customer's network. Recognizing its likely provenance, the TI company chooses to disengage and avoid researching the intrusion altogether so as to avoid rubbing Agency-A the wrong way. Six months later, Agency-A becomes aware – through public sources – that Agency-B had physically infiltrated their outfit, stolen a trove of tools, and leveraged them against institutions under Agency-A's defence purview. Was the threat intelligence producer complicit in Agency-B's stratagem by virtue of its wilful ignorance? Did Agency-A implicitly encourage this dereliction of duty?

defence needs of their customers, a 'god on the wire'-style SIGINT agency conducting fourth-party collection is in a special position to understand the context of adversary campaigns before anyone else does.

By virtue of passive or active collection to staging servers (or even operator boxes), Agency-A eliminates a huge element of uncertainty from the threat intelligence it can leverage for information assurance purposes. Rather than saying '*these are the tools Agency-B is likely to use*', Agency-A is in a position to say '*these are the tools Agency-B is currently using/getting ready to leverage*'. This high-fidelity, up-to-the-minute awareness of adversary tactics is limited only by Agency-A's visibility into Agency-B's infrastructure and is of incalculable benefit if Agency-A's internal fusion processes are efficient and effective enough[16] to leverage this information for defence while it's still actionable.

However, even if information assurance value is unlikely to be derived, other forms of adversarial learning are likely to form. As Agency-A becomes aware of the techniques leveraged by other threat actors, takes stock of the means by which Agency-B became aware of Agency-A's operations, and observes how private sector researchers latch onto other adversaries[17], it is likely to fold this knowledge into its own toolkit development and operator procedures.

Due to the scalability issues in offensive operations, most threat actors (even nation-state-sponsored actors) are not in a position to afford the luxury of becoming untraceable, truly difficult to discover, and impossible to weed out of a victim network. However, institutions like Agency-A are in a position to become the true apex predators of cyberspace and have proceeded to do so.

The following are examples of adversarial learning. The first is of mid-grade actors creating confusing overlaps that misguide campaign understanding by copying one another's TTPs[18,19]. The second is an example of a true apex predator and the techniques entailed therein that generate truly difficult – perhaps even insurmountable – problems for incident responders and threat intelligence producers going forward:

A most curious overlap

In June 2016, one of our hunting YARA rules implementing a looser heuristic for WildNeutron samples fired on an interesting trojan that we called 'Decafett'. Due to the possible overlap[20] with the notorious WildNeutron actor, we were eager to analyse it further.

This backdoor implements keylogging capabilities that monitors the victim's keyboard activity and logs it to a text file stored in the <%APPDATA%\Microsoft\Office\> directory in an encrypted form. It is also capable of launching other processes as designated by an external file named <restore.idx> located in the same directory.

Some of the strings used by the binary are encrypted with simple custom algorithms (combining XOR, SUB, ADD and other arithmetic instructions) that are different for every string. These get decrypted on the fly using inline decryption routines during program execution and are immediately re-encrypted once used.

The C&C IP address calculation is what sets Decafett apart from most other malware. It is computed based on a hostname stored in the registry or based on a default C&C hostname hard coded into the executable: download.ns360[.]info. Interestingly, the authors implemented an unusual mechanism in the backdoor to obtain the actual C&C IP address. For this, they first resolved the C&C domain through a DNS query and then XORed the IP value with the key 0x186BFB49.

The following is an example using the default hostname, download.ns360[.]info:

```
orig IP resolution: 54.251.107.25= 36 FB 6B 19 =
196BFB36
xored:  127.0.0.1   =  7F 00 00 01 = 0100007F
```

The following is an example of the default URL with an added number, download1.ns360[.]info:

```
orig:   84.45.76.100     = 54 2D 4C 64 = 644C2D54
xored:  29.214.39.124    = 1D D6 27 7C = 7C27D61D
```

This IP XOR technique was later seen implemented in other malware from the Lazarus Group. We were eventually able to attribute the Decafett malware to the Lazarus Group with greater certainty in 2016. This assessment was further confirmed when its usage was also observed during BlueNoroff[21] operations.

Another interesting feature of Decafett is its reliance on a specific and unusual dynamic DNS provider. Some of the Decafett C&C servers we identified include:

- `download.ns360[.]info`
- `update.craftx[.]biz`
- `mozilla.tftpd[.]net`
- `checkupdates.flashserv[.]net`

The top domains used for these hostnames are derived from the obscure dynamic DNS provider 'DNSdynamic' (see Figure 7).

The owner ('Eddie Davis', according to the official *Twitter* account [9]) appears to have been quite active between 2011 and 2014. The last post on his *Twitter* account (at the time of writing) is from June 2014.

We attempted to uncover other potentially malicious subdomains (in addition to Decafett's checkupdates.flashserv[.]net) on top of the flashserv[.]net

[16] This is unlikely to happen in organizations that place greater value on their collection techniques than their information assurance remit, as the possibility of revealing their level of access into Agency-B's operational infrastructure will be considered of far greater negative impact than Agency-B's intrusion and access into high-value systems. The true negative impact of over-enthusiasm with offensive techniques will likely become clear as Agency-B's access is weaponized in domains beyond cyber.

[17] *'IMHO, next-gen tradecraft will require learning from these reports and will eventually involve end-to-end decisions from development to deployment to shutdown / upgrade'* –'What did Equation do wrong, and how can we avoid doing the same?' [8]

[18] Tools, techniques, and procedures.

[19] The example of shared compromised infrastructure between ScarCruft and DarkHotel is relevant to this category as well.

[20] Further analysis ultimately indicated that there was, in fact, no code overlap with WildNeutron.

[21] It's worth noting that the Decafett malware also contains the wiping function identified by *BAE* researchers tying the Bangladesh Bank SWIFT attack to the BlueNoroff subset of the Lazarus Group, thus further tying these groups together.

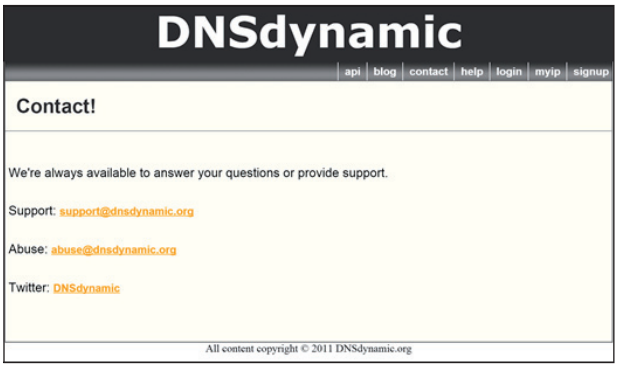

Figure 7: The top domains used for these hostnames are derived from the obscure dynamic DNS provider 'DNSdynamic'.

service offered by DNSdynamic. Interestingly, we found that the only other APT known to use this service is DarkHotel (also known as 'Tapaoux' – see Figure 8).

The DarkHotel/Tapaoux actor used these hosts in 2014, while Decafett used them in 2016. Given the long-standing adversarial relationship between these two threat actors, this indicates that the Lazarus Group likely adopted the usage of this obscure dynamic DNS provider after – and likely as a result of – DarkHotel's use.

The beginning of the end

Over the last few years, the number of apparently 'APT-related' incidents described in the media has grown significantly. For many of these, though, the designation rank of APT – an 'Advanced Persistent Threat' – is usually an exaggeration. With some notable exceptions, few of the threat actors described in the media are advanced. These exceptions represent the pinnacle of cyber espionage tools – the truly 'advanced' threat actors out there include: Equation [10], Regin [11], Duqu [12], and Careto [13]. Another exceptional espionage platform is 'ProjectSauron [14]', also known as 'Remsec' or 'Strider' (see Figure 9).

ProjectSauron is a true bleeding-edge modular cyber espionage platform in terms of technical sophistication, designed with an eye towards long-term campaigns through stealthy survival mechanisms coupled with multiple creative exfiltration methods.

The golden pupil

ProjectSauron is perhaps most impressive in its thorough address of the weaknesses and failures of other top-tier threat actors. Emulating their better features while avoiding the

```
KBLOG_ROTATE_SECS = 10800
tmp_dir = os.getenv("WINDIR") .. "\\temp\\"
drive = "C:\\"
SAURON_KBLOG_KEY = "mISfx1q2Ef/QJPO4gi6DMKD5lxeQ380knDrULcZyTF5vFNWb
create_log = function(l_1_0, l_1_1, l_1_2, l_1_3)
  local f = ""
  repeat
    w.sleep(1000)
    t1 = "b"
    t2 = "k"
    t3 = "a"
```

Figure 9: A decrypted LUA configuration script from which the name ProjectSauron was derived.

pitfalls that burned actors like Duqu, Flame [15], Regin and Equation turned ProjectSauron into such a formidable threat actor as to entirely bring into question the effectiveness of indicators of compromise (IOCs) in addressing truly advanced intrusion sets going forward.

ProjectSauron was observed emulating the following threat actors' innovations:

Duqu	Running only in memory Using compromised intranet servers as internal C&C servers Alternate encryption methods per victim Named pipes for LAN communication Malware distribution through legitimate software deployment channels
Flame	LUA scripts executed in an embedded (custom tailored) VM Secure file deletion Attacking air-gapped systems via removable devices
Regin and Equation	Usage of RC5/RC6 encryption Reliance on virtual filesystems (VFS) Attacking air-gapped systems via removable devices Hidden data storage on removable devices

Additional care was taken to avoid the following common pitfalls:

- Vulnerable or consistent C&C locations
- ISP, IP, domain, or tool reuse across different victims and campaigns
- Cryptographic algorithm and encryption key reuse

Host	IP	Country	Firstseen	Lastseen	Countseen	Tags Host IP
adoberegister.flashserv.net	124.217.251.98	MY	2014-08-23 08:57:43	2014-11-21 01:46:18	92	Tapaoux
adoberegister.flashserv.net	84.45.76.100	GB	2014-11-21 12:35:55	2016-01-15 16:09:32	490	Tapaoux
checkupdates.flashserv.net	84.45.76.100	GB	2016-01-14 08:33:55	2016-01-15 11:02:03	13	Decafett
checkupdates1.flashserv.net	84.45.76.100	GB	2016-01-14 08:35:55	2016-01-15 17:36:37	16	Decafett
javaupdate.flashserv.net	124.217.245.66	MY	2014-08-13 09:02:18	2014-11-21 05:47:52	107	Darkhotel,Tapaoux
javaupdate.flashserv.net	84.45.76.100	GB	2014-11-22 05:42:10	2016-01-13 06:00:06	301	Darkhotel,Tapaoux

Figure 8: The only other APT known to use this service is Darkhotel (also known as 'Tapaoux').

- Forensic footprint on disk
- Timestamp similarity throughout various components
- Large volumes of exfiltrated data
- Use of unusual network protocols or message formats.

Moreover, by apparently automating the modification of most (if not all) IOC-worthy elements (like mutexes, filenames, timestamps, service names, subdomains, etc.) on a per-operation basis, ProjectSauron limits the threat hunting value of happening upon a single intrusion. While carefully prepared YARA rules (the byproduct of in-depth RE) and in-memory analysis will detect ProjectSauron components, leveraging IOCs alone will not alert network administrators to their presence. This 'next-gen tradecraft'[22] makes ProjectSauron's toolkit not only truly advanced but also bespoke – tailor-made to each victim – and entails severe consequences for the future of threat intelligence[23].

Untraceable exfiltration

Another suspected feature of ProjectSauron is noteworthy in the context of the superpowers afforded to 'god on the wire'-style cyber espionage actors. As mentioned before, attempting to point out the active leveraging of passive collection capabilities on cyber operations is a hermeneutic and inexact endeavour. The following example may prove equally illustrative of a passive collection superpower, as it may simply be the result of unfortunate timing, and should be taken with a grain of salt:

One of the aforementioned innovative exfiltration methods relies on emails, which are delivered via the use of LUA scripts to a number of mailboxes allegedly controlled by the attackers. Some of these mailboxes, for instance, were hosted on free email providers such as *Gmail* or *Mail.ru*. To send emails to these mailboxes, ProjectSauron implants attempt a direct connection to the specific mail server. If that fails, it can try to deliver the emails using a 'relay' server. A deeper investigation into one of these relay set-ups revealed an unusual ProjectSauron superpower.

One of the cases we observed employed a relay set-up pointing to a machine in Chile, belonging to a massive US corporation. We contacted the owners and were allowed to check the server to better understand its configuration as an open relay. However, the server was well secured and wouldn't accept emails to mailboxes outside of its configured domains.

Two possibilities remain: despite an unlikely timeline, the server could presumably have been backdoored at some point before our investigation with special software that would intercept these emails and forward them to the attackers. Another, more fascinating, possibility is that the attackers have access to some of the fibre links between the victim's Internet space and the Chilean relay, allowing them to capture the exfiltrated information at network level.

CONCLUSION – AN UNCERTAIN FUTURE FOR THREAT INTEL

The more cautious *Twitter* talking heads rightly urge us to mince words – separating 'intrusions' from 'attacks', and 'espionage' from 'warfare'. While we agree with this cautious stance, the newfound military language of 'cyberwarfare' does shed some light on the utilitarian characteristics of 'cyber-as-a-domain' as perceived by the relevant nation-state stakeholders. Major General Amos Yadlin[24] characterizes this domain[25] as having 'unlimited range, very high speed, and [...] a very low signature'. While unlimited range and speed make cyber operations infinitely desirable to rugged military folks familiar with the true cost of conventional warfare, the final characteristic should make this domain precarious, capricious, and terrifying in its own right.

A signature weapon?

Low signature is a characteristic of weaponry whose responsible party is not immediately apparent. Mortars and improvised explosive devices (IEDs) have the have the ability to cause damage without immediately recognizable provenance. An analysis of the trajectory of the mortar or the detonation method of the IED will yield an *indication* of the parties responsible. It's interesting to note that the notion of low signature, while superficially applicable, is actually a deceptive oversimplification of the complexities posed by the use of 'cyber weaponry'.

There is a difference between a lack of identifiability, or a delay in identifiability, and the presence of identifiable traits that misidentify the perpetrators while presenting no discernible difference in quality. Where the sourcing and manufacturing processes of conventional weaponry present near-impossible-to-circumvent opportunities to identify the provenance of a weapon and its subsequent use, the infinitely reproducible nature of programs and digital tools add an insurmountable layer of complexity to identifying the perpetrators behind operations. The last bastion of the notion of definitive identifiability seems to lie with intellectual property. '*Can we identify the sole proprietor of a closed-source tool?*' In a world without perfect knowledge and immediate transference of means, the usefulness of this criterion is fading.

Attribution – beyond tool clustering and threat actor narrative interpretations – is a demand to cross the boundaries of a given domain to point to a perpetrating entity in another. Our interest is not in saying that cyber operations present no discernible signature, but rather that our analysis methods for establishing provenance beyond the boundaries of the cyber domain are limited logically by a lack of direct correlation [16] between the topology of a network, the functional organization of an institution, the position of an individual, and the ultimate symbolic value of individual targets in a geopolitical and socioeconomic order.

An added nuance directly introduced by the practice of fourth-party collection in cyber operations is the discovery and potential repurposing of closed-source tools. Where fourth-party collection will shed light on a foreign service's intelligence collection methods, recruitment tactics, and their tasking priorities, it can also yield tools, source code, procedural guidelines, identifiable information on operators and their systems, and access to IPs and known infrastructure.

[22] To borrow a Vault7 term.

[23] Further discussed in the next section.

[24] Commander of Israeli Defence Intelligence (2006–2010).

[25] In his original quote, he refers to 'cyber' as the fourth domain – an Israel-centric departure from the US-centric model that includes 'Space' as the fourth domain and displaces 'Cyber' to the fifth domain.

This final collection boon is tantamount to the means for generating misleading signatures in subsequent attacks. With strategic access to Agency-B's tools and infrastructure, Agency-A has silently violated the bounds of closed-propriety that served to identify Agency-B in the cyber domain.

While we are in no way advocating the unhealthy scepticism that arises from viewing every campaign as a potential false flag and every attribution claim as a self-serving misstep, we'd rather point to the extreme outlier of fourth-party collection-enabled misattribution and signature falsification as a statement indicative of the logical boundaries of the knowledge that can be derived from single-source investigations into perpetrating actors in cyberspace.

Treacherous surroundings for threat hunters

As the subject of study becomes more visibly complicated, the defenders involved are facing their own problems. The honeymoon with threat intelligence is coming to an end. While the 'APT' moniker has served the diabolical needs of RSA showfloor salesfolks, threat intelligence has not always proven a stable source of revenue for companies young and old. As behemoths reconsider their threat intelligence investments, a series of complications tormenting the immediate future of threat intelligence come into view:

- Regional balkanization of research capabilities and lack of defence of the academic and objective value of threat intelligence research beyond PR and marketing.
- Greater adoption of non-telemetry-friendly operating systems.
- Rejection of anti-malware tools as 'overly intrusive' promotes increased adoption of userland EDR-style tools incapable of detecting sophisticated threats.
- Lack of adequate telemetry generation and aggregation becomes the norm.
- 'Next-Gen' or 'machine-learning' tools focus on specific processes (and not the whole OS) to avoid overhead while claiming full protection.
- Reliance on scraping multiscanners[26] to 'supplement' detections gives a false sense of ubiquitous coverage.
- 'Modern' solutions and alternative OSs[27] avoiding complex heuristics and process tracing to limit processing-power overhead are recreating blindspots already addressed by the anti-malware industry.
- Two years after providing in-the-wild proof of abuse of firmware by the Equation Group, firmware is no more accessible for analysis than it was then.
- While memory forensics capabilities have increased in leaps and bounds[28], a concern with system stability has kept most modern anti-malware solutions from attempting greater strides to address memory-resident malware.

These conditions, presented in no particular order, blend together to outline the treacherous surroundings of a nascent industry whose practitioners are perhaps too fascinated with the intricacies of their study to see the earth fragmenting beneath their feet. In the cyber-arms race, advanced threat actors are pulling away at an alarming speed. While we have become better at spotting them and communicating this knowledge for other defenders to take measures en masse, we should be troubled by the tendency towards crippling telemetry generation and neglecting the advancement and ubiquity of heuristic capabilities.

It's important to highlight two important gains provided by threat intelligence that are often taken for granted:

First, while there is a great community of vulnerability researchers out there providing insights to improve the code relied on by all users regularly, threat intelligence researchers have been providing increasing insights into exploits being leveraged in the wild with proven malicious intent. While most threat intelligence production may be tied to a particular company, intended to protect a specific subset of consumers, by reporting these zero-days to the relevant codebase maintainers, the benefit is provided for all users in a given ecosystem regardless of their chosen security software provider. *What happens to this important function as threat intelligence production is further disincentivized?*

Secondly, while the private sector is largely motivated by big, flashy announcements, the resulting byproduct is tantamount to a series of stories that collectively amount to a public understanding of fifth-domain operations. This body of literature is the only relevant information when evaluating policy to regulate cyber measures and norms[29]. It is largely reliant on private, non-reproducible data sources that may very well fade away and disappear as threat intelligence companies are dissolved or acquired. Subsequent private reports and proprietary data that never entered the public domain to begin with would actually disappear from within these companies without relevant maintainers. The question that should burn within us at the sight of these developments is: '*What will happen to this Homeric history of human adversarial incursion into the fifth domain in their absence?*'

An unhealthy obsession

Finally, we'd do well to reiterate a disdain for the obsession with cross-domain attribution based on single-domain information and analysis. No element of cyber threat intelligence has incited more ardent ignorant opinionated ineffective bickering than whether a given attribution claim is right/wrong/justified/believable or acceptable given someone else's gut feeling of what an unknown actor might find reasonable to do on a given day. No element of this cacophonous symphony played on 140-character instruments speaks to the true nature and limitations of our practice.

Attempts at armchair attribution are instances of running up to (and past) a methodological cliff into an area that is simply beyond private sector fifth-domain capabilities. It's within reach for all-source agencies, it's within reach for Agency-A style institutions, by virtue of their visibility simultaneously *within and **beyond*** the fifth domain. For everyone else, one is tempted to (mis-)apply Ludwig Wittgenstein's Tractatus

[26] It is interesting to consider a future where these vendors win majority market share: based on whom will they mimic their detections then?

[27] A lack of complex heuristic monitoring solutions in *MacOS* and *Linux*.

[28] Thanks in large part to the Volatily Project [17].

[29] Although think tanks and lawyers somehow manage to ignore it all the same, in favour of hyperbolic and technically impossible overstretched analogies.

Logico-Philosophicus in its final admonition[30] of hyperextending beyond logical boundaries to make statements of logical fact. But in reality, while it's the logical nature of the fifth domain that limits what statements we can reasonably make in observation of these advanced campaigns, it's a desire for relevance and the limelight that is ultimately diminishing the value of the nascent and much needed practice of threat intelligence currently in need of its own ardent defence.

REFERENCES

[1] http://www.spiegel.de/media/media-35684.pdf.

[2] CSEC SIGINT Cyber Discovery: Summary of the current effort. REPLICANT FARM output, pp.10–11. https://edwardsnowden.com/wp-content/uploads/2015/01/media-35665.pdf

[3] https://securelist.com/energetic-bear-more-like-a-crouching-yeti/65240/

[4] DarkHotel Attacks in 2015– https://securelist.com/darkhotels-attacks-in-2015/71713/.

[5] Friedrich-Ebert-Stiftung (FES). http://www.fes-caucasus.org/.

[6] https://github.com/hackedteam/rcs-db/tree/master/scripts.

[7] Re: Nomi agente per sistema di test. https://wikileaks.org/hackingteam/emails/emailid/106487.

[8] https://wikileaks.org/ciav7p1/cms/page_14588809.html.

[9] https://twitter.com/DNSdynamic.

[10] https://securelist.com/blog/research/68750/equation-the-death-star-of-malware-galaxy/.

[11] https://securelist.com/blog/research/67741/regin-nation-state-ownage-of-gsm-networks/.

[12] https://securelist.com/blog/incidents/32463/duqu-faq-33/.

[13] https://securelist.com/blog/research/58254/the-caretomask-apt-frequently-asked-questions/.

[14] https://securelist.com/faq-the-projectsauron-apt/75533/.

[15] https://securelist.com/the-flame-questions-and-answers-51/34344.

[16] 'TREASUREMAP', NSA project to map entities across domains. https://snowdenarchive.cjfe.org/greenstone/collect/snowden1/index/assoc/HASH01c8/6ed8dd75.dir/doc.pdf.

[17] http://www.volatilityfoundation.org.

[30] 'Whereof one cannot speak, thereof one must be silent'. Proposition §7, *Tractatus Logico Philosophicus*, Ludwig Wittgenstein, 1921.