Vladimir Tivanski

Professor Rishab Nithyanand

CS 3640-01

1 September 2022

Technical Analysis of NSA's PRISM and MUSCULAR Surveillance Programs

Recently, citizens of the United States of America have become well acquainted with digital technologies. Now more than ever a vast majority of Americans use digital technologies in their day to day lives. According to a study conducted by the Pew Research Center, 97% of Americans today own a cellphone, including 85% who own a smartphone, up 50% from the 35% of Americans that owned smartphones in 2011. Everyday the American population relies heavily on communication technologies to contact others for various private reasons. Many people would think it safe to assume their privacy when communicating over these technologies, but in reality the National Security Agency(NSA) actively monitors the calls of hundreds of millions of unknowing Americans.

Since the terrorist attacks of September 11, 2001, the NSA has significantly expanded their mass surveillance program. Disclosures have shown that the NSA continues to actively monitor Americans' communication activities including international calls, text messages, emails, and web activity.

The NSA's unrivaled authority to oversee such communications boils largely down to a law passed by congress called the FISA Amendments Act of 2008(FAA). "The FISA Amendments Act of 2008 (FAA) gives the NSA almost unchecked power to monitor Americans' international phone calls, text messages, and emails — under the guise of targeting foreigners abroad." (ACLU) In 2013, former NSA contractor and whistleblower Edward Snowden

confirmed the immense scale of the NSA's control on international communications, and many recent disclosures additionally show the NSA actively monitors purely domestic communications as well. An article by the American Civil Liberties Union well summarizes the impact of the FAA. "The rules that supposedly protect Americans' privacy are weak and riddled with exceptions." (ACLU)

Executive Order 12333 (EO 12333) signed into legislation by President Reagan in 1981 is a document that grants the NSA and other intelligence agencies complete authority to collect foriegn intelligence outside the United States. The NSA actively uses this bill to justify the collection of American information. For example, the NSA collects the location information of billions of phone calls every day and records all phone calls into and out of the United States. Moreover, the NSA uses this legislation to justify intercepting data from Google and Yahoo as that information travels between data centers located outside the United States. According to an interview with Edward Snowden, "The National Security Agency has secretly broken into the main communications links that connect Yahoo and Google data centers around the world, [...] By tapping those links, the agency has positioned itself to collect at will from hundreds of millions of user accounts, many of them belonging to Americans. The NSA does not keep everything it collects, but it keeps a lot." (Washington Post)

According to a top-secret report from January 9, 2013, the NSA sends millions of records collected from breached Yahoo and Google networks to data warehouses at the NSA's headquarters. "In the preceding 30 days, the report said, field collectors had processed and sent back 181,280,466 new records — including "metadata," which would indicate who sent or received e-mails and when, as well as content such as text, audio and video." (The Washington

Post) Clearly the NSA is able to gather vast amounts of user and communication data, but how specifically are they able to get away with this without alerting the public?

One way the NSA accomplishes this is via direct access to the servers of providers that provide services to facilitate communication. A program code named PRISM claims to collect data directly from the servers of major US service providers. In 2013, a 41 slide presentation detailing PRISM was leaked to the public. In the slideshow, companies like Microsoft, Yahoo, Google, Facebook, PalTalk, YouTube, Skype, AOL, Apple, and (with "plan to add") Dropbox are all listed as data providers.

Despite the leaked presentation, many of these tech giants deny ever cooperating with the NSA. 'We have never heard of PRISM," said Steve Dowling, a spokesman for Apple. "We do not provide any government agency with direct access to our servers, and any government agency requesting customer data must get a court order."' (The Washington Post) Like Apple, many of the other companies also issued statements refuting the document and claiming that they only provide access to user data when a legally binding order is received. Regardless of these claims, with PRISM the NSA clearly has access to all of these companies' servers and all the communication data that they store and process. According to the document, PRISM is prided as "one of the most valuable, unique and productive accesses for NSA" and is titled "*The SIGAD Used **Most** in NSA Reporting*" (NSA)

The disclosed report additionally details what data specifically is collected, what process is used to parse data, and how that data is used after its collection. The document states that what specific data is collected varies from provider to provider, but generally, e-mails, chats (voice and video), videos, photos, stored data, voice over internet protocol(VoIP), file transfers, video

conferencing, notifications of target activity (logins etc.), and social networking details are all subject to the collection of PRISM.

The process as to which PRISM goes about collecting this data is outlined in a chart titled PRISM Tasking Process. The first step of the chart is for a "Target Analyst" to input "selectors" (search query) into the "Unified Tasking Tool(UTT)." Leaked instructional screen captures show that the program is a web-based application which formats a search query for a specific user. The target analyst has options to specify the target's name (which can be queried from a database "NYMROD"), type (which is shown to be "person" by default implying that it is possible to query something other than "person"), nationality, location, and classification. Additionally the target analyst is required to enter a short justification, which the NSA strictly outlines to be "just one sentence" and demanded to be kept as vague as possible and a foreignness factor, which is a reason to believe the target is outside the United States (which is a requirement for legal usage of the program).

After this step the diagram shows a two way part in the task process. The part outlines two functions of the UTT. On the left the diagram outlines the task process for surveillance tasks(involving the use of provider data) and on the right the diagram outlines the task process for searching stored communications(internally in government databases).

In order for a surveillance task to proceed it must be approved twice. The first verification is done by a FAA Adjudicator (S2). According to the leaked document such an adjudicator exists for each of the NSA's "Product Lines", which are the NSA's departments for various specific issues like counter terrorism. The second check must then be conducted by NSA unit 343 for Targeting and Mission Management. After these two checks, the request is unlocked in the Unified Tasking Tool and further progress is allowed. In contrast, for searching stored

communications, the first validation is conducted by the Special FISA Oversight and Processing unit (SV4), and the second check is again completed by NSA unit 343.

After getting verified a task is then able to proceed from the Unified Tasking Tool to a Site Selector Distribution Manager named PRINTURA. The actual data collecting process once a target is identified is not conducted by the NSA but rather by the FBI. After the NSA finds a target it wants to surveil it passes said target to the FBI's Data Intercept Technology Unit(DITU). At first DITU managed the FBI's internet monitoring programs Omnivore and Carnivore, which gathered raw data by tapping into various ISP locations then parsed the data using the Packeteer en Coolminer tools. The FBI likely still uses these tools but it is also described in the leaked document that they have direct connections with online service providers.

After raw data is gathered by the FBI, its storage is then outlined in the document through the "Collection Dataflow". Once the data is collected it is sent back to the PRINTURA system which The Washington Post describes as a program that automates traffic flow. Essentially a system named SCISSORS and a system labeled Protocol Exploitation conducted by NSA units T132 and S3132 respectively sort the data across the NSA databases TRAFFICTHIEF(for metadata about specific email addresses), MARIANA(for internet metadata), MAINWAY(for telephone and internet data contact chaining), NUCLEON(for voice communications), and PINWALE(for video content, FAA partitions, and Digital Network Intelligence(DNI) content, which is internet content like forum postings or chat messages).

PRISM is just one of the NSA's tools for collecting people's communication data (collectively labeled SIGADs by the NSA). Despite having lawful access to user accounts through PRISM the NSA jointly with the British GCHQ additionally operate a SIGAD codenamed MUSCULAR. "From undisclosed interception points, the NSA and the GCHQ are

copying entire data flows across fiber-optic cables that carry information among the data centers of the Silicon Valley giants." (The Washington Post) Essentially, the NSA and the GCHQ are tapping into fiber-optic cables to steal data as it flows through the cable.

Massive tech giants, like Google, have "fortresslike" data centers across multiple continents that are connected with thousands of miles of fiber-optic cable to protect against data-loss and system slowdowns. Since these data centers and cables are outside United States boundaries the NSA is able to abuse a legal loophole to steal this data. "Outside U.S. territory, statutory restrictions on surveillance seldom apply and the FISC has no jurisdiction." (The Washington Post) Intercepting communications outside the United States allows the NSA to legally perform large scale data seizure since the NSA is allowed to presume that anyone using a foreign data link is a foreigner.

In an NSA presentation slide titled "Google Cloud Exploitation" the NSA outlined a security flaw in Google's dataflow. As data is transferred from the "Public Internet" to Google's internal "Google Cloud", it travels over a link where the data is added and removed meaning for some time it is unprotected. MUSCULAR operates in this area. MUSCULARS first step is to load all the flowing data into a buffer that can hold several days worth of traffic before recycling it for storage space. From this buffer the NSA decodes the data formats that the companies use in their clouds. After this step the NSA can proceed to collect the data in a similar process as PRISM. Notice how all the verification steps are bypassed in this process.

Senator Ron Wyden well summarizes the significance of this "foreign communications" loophole. "'Thirty-five years ago, different countries had their own telecommunications infrastructure, so the division between foreign and domestic collection was clear,' Sen. Ron Wyden (D-Ore.), a member of the intelligence panel, said in an interview. 'Today there's a global

communications infrastructure, so there's a greater risk of collecting on Americans when the NSA collects overseas.'" (The Washington Post) Nowadays, all major tech giants transfer data overseas. The NSA's restriction to operate on foriegn communications to "protect" Americans' privacies is a red herring. Since communications from all around the world are traveling all around the world there is no way to tell the difference between an American communication and a foreign communication without looking at the communication itself.

Moreover, the FISC court which oversees government's surveillance programs operates in total secrecy. When the court was first established in 1978 its primary purpose was to oversee individual surveillance to determine whether a target was foreign. Now however, the court rules over mass surveillance programs and assesses their constitutionality, all out of the public eye. Without public interpretation on constitutionality, who is there to check the court's power? The fourth amendment of the United States constitution protects the people from unreasonable searches and seizures by the government, but legal loopholes allow government agencies to dishonestly bypass this amendment and impede on the privacies of American citizens. Despite claims from tech giants like Microsoft whose slogan was "Your privacy is our priority", the NSA is capable of monitoring the data of nearly all Americans who use the internet.

Work Cited

https://www.pewresearch.org/internet/fact-sheet/mobile/

I used this research publication that studied the ownership and demographics of mobile phone ownership over time to gauge the extent of people that the NSA's surveillance program could apply to.

https://www.aclu.org/issues/national-security/privacy-and-surveillance/nsa-surveillance

I used this general information article about NSA surveillance and related legislation as a general exposition to NSA's surveillance program and the laws that apply to these operations. This article also introduced me to the legal loopholes that government agencies use to collect more communications data.

https://www.washingtonpost.com/world/national-security/nsa-infiltrates-links-to-yahoo-google-data-centers-worldwide-snowden-documents-say/2013/10/30/e51d661e-4166-11e3-8b74-d89d714ca4dd_story.html

I used this article from The Washington Post as a second step from the ACLU article to gain a more detailed perspective about which technologies specifically are used by the NSA to collect data.

https://www.theguardian.com/world/2013/jun/06/us-tech-giants-nsa-data

I used this article from The Guardian as I was reading the leaked PRISM slides to ease my understanding of the slides and learn more about their general context and reactions to the slides. I also used this article to learn about the MUSCULAR SIGAD.

https://en.wikipedia.org/wiki/File:Target-analyst-rationale-instructions-final.pdf

I used this document to study the Unified Tasking Tool mentioned in the disclosed PRISM slides.

https://www.washingtonpost.com/wp-srv/special/politics/prism-collection-documents/

I used this Washington Post article to aid my research of the disclosed PRISM slides.

https://www.washingtonpost.com/investigations/us-intelligence-mining-data-from-nine-us-internet-companies-in-broad-secret-program/2013/06/06/3a0c0da8-cebf-11e2-8845-d970ccb04497_story.html?tid=a_inl_manual

This article from The Washington Post is very similar to the article I read from The Guardian giving context to the PRISM slides. I used this article to affirm the information that The Guardian was presenting and as a second source for information about the PRISM slides.

https://nsa.gov1.info/dni/prism.html

This is the leaked document from the NSA which outlines what providers collaborate with PRISM, what data is collected from these providers, what process is used to collect this data, and what is done with this data after collection. I used this raw source to base my technical analysis of PRISM.

https://www.electrospaces.net/2013/07/new-insights-into-prism-program.html

This article significantly helped me decipher the diagrams on the disclosed PRISM slides. This article additionally helped me understand different branches of the NSA the coordinate to make PRISM work.

https://www.law.cornell.edu/constitution/fourth_amendment

I used this Cornell article to learn about the fourth amendment of the Constitution.

https://www.washingtonpost.com/world/national-security/nsa-broke-privacy-rules-thousands-of-times-per-year-audit-finds/2013/08/15/3310e554-05ca-11e3-a07f-49ddc7417125_story.html

I used this Washington Post article to gain some exposure to NSA security breaches.