# Mass Surveillance and Metadata Collection in the US

Stephanie Shin

ENGR 188EW

Fall 2020

Mass surveillance is a new and growing technology that affects all individuals whether they are aware of it or not. As technology makes it easy for people to stay connected from the comfort of their smartphones and computers, it has become an integral part of everyday life. But with this new normal of staying connected online comes the question of whether or not our online lives are truly private, and what it means for the government to have our personal information. Edward Snowden is a famous whistleblower who revealed the NSA's mass surveillance programs to the public, which created an uproar of criticism questioning what the US government was doing with the information it was collecting (Cayford and Pieters 2018). It was revealed that the NSA program, PRISM, was targeting both non-US and domestic internet communications for metadata collection, which sparked global concern (Landau 2013). Many proponents argued alongside the U.S. government, stating that mass surveillance increases security and serves as an efficient method for preventing domestic and foreign terrorism (Cayford and Pieters 2018). On the other hand, opponents claim that these programs reflect a breach of the privacy of users' data and the government's lack of upholding the fourth amendment. This right to privacy protects citizens from excessive governmental power, as upon acquiring metadata about millions of users, the government is given the option to interfere with its citizens, whether or not it decides to take that course (Stahl 2016). So when considering the effectiveness of these mass surveillance programs, it is important to consider the cost vs. benefit of running the program, privacy rights of internet users, as well as the government's duty to protect its citizens. Therefore, mass surveillance should be utilized to monitor large populations and filter out data on suspicious individuals. But in doing so, the U.S. government should impose strict laws on intelligence collection, and not store any metadata on the rest of the population (both domestic and foreign) to protect the privacy of those people.

Mass surveillance has largely come into practice after 2001, as following the 9/11 terrorist attack, "the criteria to conduct surveillance, whether electronic or otherwise, loosened" (Landau 2013). After the 9/11 attack, the government reconsidered its prioritizations regarding

different types of communications, now deeming American phone metadata as both important and legal to collect (Cayford and Pieters 2018). President Bush also increased the NSA's authority to "warrantlessly wiretaps international communications", which further relaxed international data collection laws that had already been quite loose (Landau 2013). Snowden's report reflects the effects of these loosened laws, revealing that IP addresses were being stored as well as telephone metadata to perform "contact chaining", which stores information about numbers called by and to certain targets (Landau 2013). This information proved to be alarming, as the US government had been lying to the public about these types of surveillance programs, as before the leaks, government officials denied claims of these practices while testifying in Congress (Landau 2013). Not only does this raise questions about the legality and legitimacy of mass surveillance, but it also brings to light the concept of strategic intelligence, which is a broad version of intelligence that collects information without knowledge of what type of information will be found (Cayford and Pieters 2018). This is much different from targeted surveillance, where individuals (rather than masses) are monitored when there is suspicion and reason to do so (Cayford and Pieters 2018). Consequently, strategic intelligence is reliant on relatively new technologies that allow for mass data collection and efficient analysis of what is collected.

This new large scale surveillance is only possible with the technological advancements made within the last several decades. Although the collection of phone/internet data can be accomplished using a variety of technologies, we will mainly discuss the use of cable splitters to obtain that data. Fiber-optic cables enable global communication by carrying internet traffic through electrical signals (Cayford, van Gulijk and van Gelder 2014). These fiber-optic cables can be tapped through the consent of telecommunications companies to install wire splitters at various locations (Cayford, van Gulijk and van Gelder 2014). Wire splitters work by receiving a single input signal and splitting the signal through two different outputs (Cayford, van Gulijk and van Gelder 2014). Although the resulting signals are weaker, both the original destination and the NSA can get the same copies of information (Cayford, van Gulijk and van Gelder 2014). This data can then be sent to a new location to either be stored or processed. Deep Packet Inspection (DPI) technologies can analyze this internet traffic extremely quickly, extracting the necessary information from IP addresses to actual content (Cayford, van Gulijk and van Gelder 2014). The NSA made use of TURMOIL to collect data, and XKEYSCORE to store "3 days worth of raw packet data and 30 days worth of metadata in the local caches" (Cayford, van

Gulijk and van Gelder 2014). Further, XKEYSCORE can also process the data when a query is created with either a hard selector such as an email, or a soft selector such as a broader category of information (Cayford, van Gulijk and van Gelder 2014). Based on this query, the selected information is then sent back to the NSA's database for further investigation (Cayford, van Gulijk and van Gelder 2014). XKEYSCORE proves to be a powerful tool as there are around 150 XKEYSCORE sites around the world along with splitters in most major cities in the U.S., allowing the NSA to intercept "99% of U.S.-only traceroutes" (Cayford, van Gulijk and van Gelder 2014). These technologies reflect mass surveillance as the information carried by internet traffic is both copied and stored in large quantities.

Mass surveillance can be deemed effective and useful if it increases security and provides a larger benefit than the cost of running it. However, it is difficult to measure the effectiveness of a program designed as a preventative measure. Mass surveillance is intended to detect and prevent terrorism, which makes it hard to organize its successes into quantifiable data. The NSA director cited that 54 attacks had been thwarted from the information collected from mass surveillance programs (Cayford and Pieters 2018). We can further imply that through stopping these terrorist attacks, the technology was able to save lives, infrastructures, etc. which can also be argued as a great success of mass surveillance. Further, we can analyze the effectiveness of mass surveillance in the U.K., where their Home Secretary attributed 95% of their criminal cases to utilizing communications data (Cayford and Pieters 2018). But despite minimal evidence given by the NSA of direct positive results from this technology, it is also argued that security measures that simply give the public a sense of protection (regardless of its true effectiveness), can deter terrorist activity which in effect does create greater security (Cayford and Pieters 2018). Despite these numbers, the NSA continues to be vague in their findings and reports of their agency in general rather than of the specific mass surveillance technologies (Cayford and Pieters 2018). This can make it difficult to quantify whether or not the mass surveillance programs are truly integrated into the effectiveness of providing security to the public.

Although there are examples of mass surveillance resulting in greater security, there are also many concerns that come with this technology. The NSA primarily argues for the usage of surveillance technology by citing the number of terrorist attacks intercepted from interfering with communications data. However, using risk assessment tests, the cost of spending on anti-terrorism measures largely outweighed the benefits, and increased funding does not show

noticeable changes in progress (Cayford and Pieters 2018). Terrorist attacks also tend to be extremely uncommon, and most security measures are implemented to provide a sense of security, whether or not they are truly effective (Cayford and Pieters 2018). Oftentimes, many terrorist plots are also exaggerated or much less serious than expected, which further contributes to the idea that terrorism prevention is far too costly for the small benefits it reaps (Landau 2013). Further, the Bush administration's wiretapping efforts were reported to have a small role in the FBI's counterterrorism efforts upon investigation (Landau 2013). It is also difficult to truly isolate the effects of these specific surveillance technologies from all of the agency's resources combined as a whole.

Aside from the cost-benefit analysis of mass surveillance technology, there is a large issue regarding internet/phone users and their right to privacy. There are not many laws in the U.S. protecting metadata even though it provides some of the most detailed content concerning users' data (Landau 2013). It is not given the same Constitutional protections as actual data content since it involves third parties (telephone companies) (Landau 2013). The government can easily obtain metadata without probable cause, which questions whether or not the government will abuse this power (Landau 2013). But it is also notable that simply having the power to obtain metadata can already be an issue, regardless of whether or not that power is truly exploited. According to the fourth amendment of the constitution, U.S. persons are protected against unreasonable searches and seizures, which can be directly applied to mass surveillance (Cayford and Pieters 2018). Al Gore states that the NSA surveillance does violate the Constitution as "it isn't acceptable to have a secret interpretation of a law that goes far beyond any reasonable reading… and then classify as top secret what the actual law is" (Landau 2013). If a governmental agency can look at the metadata of citizens without a strong basis, it can be interpreted as a violation of the right to privacy. Although there are various ways to interpret the constitution, potentially violating the fourth amendment and purposefully hiding information about the practice from the public will likely cause political turmoil. Mass surveillance also causes discourse between the U.S. government and foreign nations. Because the U.S. has much looser laws when it comes to foreign data collection, there are concerns from European leaders about the invasion of privacy of their citizens (Landau 2013). Although there are separate laws for the EU which have established privacy protection for EU citizens, the NSA's collection of data on these nations undermines the protection the EU had guaranteed for its people (Landau

2013). It is further necessary to create a distinction about whether the mere collection of data is a threat to privacy, or the analysis and storage of it is. Privacy can be defined as "the individual's right to exclude others (such as the government) from access to certain kinds of information" (Stahl 2016). By allowing the government access to users' data, which for many years was occurring without consent, it can be said that this is a direct violation of those people's right to privacy over their information. With indiscriminate mass surveillance, the government is essentially given control over information that most users never surrendered in the first place (Stahl 2016). If the simple access to data gives the government more control and influence over its nation, we can argue that the mere collection and storage of metadata is unethical. But it is also good to note that while this storage of data is oftentimes happening unbeknownst to users, it can be seen as ethical as long as the information is not being misused.

There are many pre-existing laws and regulations in place to help protect the privacy of U.S. citizens. For example, "Title III of the 1968 Omnibus Crime Control and Safe Streets Act and the 1978 Foreign Intelligence Surveillance Act (FISA)" require a strong and probably cause for wiretapping, such as high likeliness of a suspect committing a crime or the suspect being of a foreign power (Landau 2013). However, many laws were implemented before the modern age of technology and 9/11, and many lawmakers likely did not intend for technology to become such a massive part of daily life. Furthermore, many laws have been vaguely interpreted to justify controversial actions. "Section 215 of the 2001 USA Patriot Act authorizes the collection of business records" which typically refers to driver's license records, credit card records, etc. (Landau 2013). But the government has "justified requests for domestic telephone metadata delivered in bulk, not individualized requests" which allows mass surveillance to occur through the vague interpretations of the law (Landau 2013). Therefore, it is imperative that the U.S. institutes strict laws and consequences for agencies who decide to participate in mass surveillance to ensure the process is highly monitored to warrant the safety of people's privacy.

To preserve internet users' privacy while allowing the U.S. to protect its citizens, the U.S. should make use of mass surveillance under strict laws and regulations while being transparent about its practices. Through Snowden, it was clear that the U.S. was not always truthful about its operations, which creates distrust between citizens and their governing powers. In this type of democracy, the public needs to trust these government institutions, or else these institutions will suffer in performance and ultimately fail at their goal of serving the public (Cayford and Pieters

2018). Because of the sheer power and influence the government has over its nation, we must be conscious about how that power can be misused given these huge quantities of data. Therefore, the government should not store metadata unless necessary. For instance, there must be quantifiable evidence or a high probability of suspicion on a target to store that metadata, while there does not necessarily need to be as strict of a rule on the mere collection of it. They should also provide evidence of the effectiveness of these mass surveillance programs to ensure that funding is being appropriately allocated toward these programs. However, it must be noted that this may not be completely accomplishable, as the government also cannot be completely transparent with the public in its operations. A lot of practices are kept fairly secretive to ensure its effectiveness and prevent enemies from undermining those procedures. Therefore, government agencies should be trusted to keep imperative details private, while still being able to provide enough evidence that their work is producing output to ensure greater security.

Mass surveillance is a powerful tool that must be monitored to ensure its ethical use. It is an obvious example of how evolving technology does require new laws and regulations to be fairly integrated within society. It can have enormous implications within the U.S., and can even affect nations abroad as well. However, those using these technologies must uphold people's rights to privacy, and be transparent about what procedures are being performed. This allows the government to perform its duty in protecting and serving while respecting the individual's right to privacy.

References

Cayford M, Pieters W. 2018. The effectiveness of surveillance technology: What intelligence officials are saying. The Information Society 34:88-103. [accessed 2020 Oct 30]

Cayford M, van Gulijk C, van Gelder P. 2014. All swept up: An initial classification of NSA surveillance technology. Safety and Reliability: Methodology and Applications:643-650. [accessed 2020 Oct 30]

Landau S. 2013. Making Sense from Snowden: What's Significant in the NSA Surveillance Revelations. IEEE Security & Privacy 11:54-63. [accessed 2020 Oct 30]

Stahl T. 2016. Indiscriminate mass surveillance and the public sphere. Ethics and Information Technology 18:33-39. [accessed 2020 Oct 30]