China's Data Collection on US Citizens: Implications, Risks, and Solutions

Ming Shin Chen
Georgia Institute of Technology, Ivan Allen College of Liberal Arts, North Ave NW, Atlanta, GA 30332
Corresponding author: mingschen@outlook.com
Keywords: OPM hack; Marriott Starwood; China; cybersecurity; counterintelligence

Executive Summary: The People's Liberation Army of China has been linked to several major data breaches targeting the personal data of American citizens, including the hacks on the Office of Personnel Management (OPM), Marriott Starwood, United Airlines, and Anthem Health Insurance, amongst others. These data breaches include personally identifiable information on millions of American citizens, including full names, Social Security Numbers (SSNs), job and income data, passport numbers, and flight histories. The data breaches also included the loss of roughly 18 million copies of Standard Form 86, which included personal data, including individuals' past substance abuse, gambling habits, and history of psychiatric care (Koerner 2016). The fact that the cyber intruders did not target financially valuable data, coupled with the long duration of these cyber espionage campaigns, indicate the involvement of a state-backed actor. Several post-breach investigations conducted by cybersecurity firms including ThreatConnect, and Mandiant, in addition to investigations undertaken by the US government, have attributed the attacks to a Chinese state-backed actor (Armerding 2016; Mandiant; Threat Connect 2015). It is believed that the information gathered from these data breaches is being compiled into a database by intelligence services in China, who seek to target US citizens for intelligence gathering purposes. Citing evidence from the goals and operations of Chinese intelligence services, this report makes the case that Chinese intelligence services will use this database to identify, target, and recruit US informants.

The report finds that Chinese intelligence services, namely the People's Liberation Army (PLA) and the Ministry of State Security (MSS), were complicit in the creation and use of this database. While the PLA conducts the bulk of the cyber offensive operations to collect information for the database, the MSS, China's premier foreign intelligence agency, is likely to make use of the database. Based on the operating goals of the MSS, it is likely the database will be used to aid in the agency's informant recruitment process. The MSS's informant recruitment process often begins with virtual communications and ends with actual "recruitment" occurring in mainland China. The report found that the MSS follows 5 key steps in its informant recruitment process, including (1) "spotting"; (2) "assessing"; (3) "developing"; (4) "recruiting"; and (5) "handling".

To counter the threat posed by Chinese intelligence services, this report seeks to identify high-value strategic targets which would contribute greatly to the database's utility in recruiting US informants, following the MSS's five-step informant recruitment process. The report further sought to devise countermeasures to protect these strategic targets, including tighter cybersecurity standards, data privacy regulations, and counterintelligence efforts. Key targets identified include:

(1) Data broker companies, specifically those that gather "people" data. This type of data includes information like names, contact info, SSN, education, and job information, which could be used to

"spot", or identify American citizens of interest, the first step in the informant recruitment process. The report recommends enacting federal regulations on the data collection practices and cybersecurity standards of data broker companies, maximizing cyber defenses while minimizing data exposure

(2) Open-source social media platforms like LinkedIn, which may be used to identify and target US citizens. This data will be used to "assess" and "develop" potential informants. Several reports from Western intelligence agencies revealed that Chinese intelligence sources have utilized LinkedIn to reach out to potential informants, posing as headhunters with the appeal of career-advancing opportunities. The report recommends US counterintelligence services coordinate efforts with LinkedIn to identify, publicize, and remove the accounts of the fake headhunters operating on the social media site.

(3) The Department of Homeland Security's Flight Tracker stores data on passport numbers and the arrival and departure flight history for individuals' dating five years back. This data could be used in the "recruitment" step of the informant recruitment process, as evidence from the MSS's operations indicated physical recruitment encounter frequently occur in mainland China. The information could be cross-referenced from flight histories from the United Airlines hack, and passport numbers from the Marriott Starwood hack. Due to the value its database would contribute to already stolen stores of personal data, this report issues an advisory warning to the DHS. The agency should work to bolster its cyber defense infrastructure, in addition to efforts to detect malicious intruders in the database.

(4) This report revealed the extent to which Chinese intelligence services are working to gather human intelligence in the US, and the ways in which the personal data collected on US citizens might be used to help them in this process. While none of the data stolen in the OPM, Marriott Starwood, and United Airlines are known to have turned up on the dark web, this report finds that these data breaches present a significant national security risk to the United States and its citizens.

I. Introduction: China's computer network operations against the US:

Cyber-attacks targeting the personally identifiable information of US citizens have occurred in different ways, taking different attack vectors and targeting a variety of data types. Nonetheless, the characteristics of the stolen data indicate that the Chinese government perpetrated these attacks. Understanding the nature of these attacks will provide a basis to evaluate the goals and future targets of the Chinese state.

In April 2015, a security engineer at OPM detected unusual outbound traffic while conducting routine maintenance of the agency's digital network. The unexpected signal pinged a site called opmsecurity.org, which the engineer did not recognize as one of the official domains of OPM. The OPM network was being breached, but it was unclear by whom or for how long. When the domain name was traced to the pseudonym Steve Rogers, an Avenger from the Marvel superhero universe, PLA Group 61398 became the prime suspect in the breach of the agency's database. The ode to the Avengers superhero was recognized as a trademark of the shadow-hacker group, which was also responsible for the hack of the health insurance company Anthem a few months prior. PLA Unit 61398, a state-sponsored Advanced Persistent Threat (APT), has been known to use the cyber offensive to advance political, economic, and military objectives. However, as the group generally conducts industrial and economic espionage, the motive for the Anthem and OPM hacks became less clear.

Ultimately, the data breach on OPM compromised over 4 million federal employees' information. OPM's digital archives contain roughly 18 million copies of Standard Form 86, a 127-page questionnaire for federal security clearance that includes personal information including Social

Security Numbers (SSNs); residency and educational history; employment history; information about immediate family and other personal and business acquaintances; health, criminal and financial history; personal background information, coupled with sensitive information including applicants' substance abuse; gambling habits; and psychiatric care. The hackers gained access to the complete personnel files of 4.2 million employees, past and present, including 5.6 million government employee fingerprints. The data compromised dates back to 1985, though most of the data that was targeted was from the year 2000 onwards.

Several data breaches that have been attributed to the Chinese government have been discovered since. In 2018, two Chinese hackers were indicted for their role in hacking into the US Navy Personnel files, stealing personal data on more than 100,000 US Navy personnel (Nakashima et al. 2018). The two hackers, thought to be working for the Ministry of State Security (MSS), stole data including names, SSNs, date of birth, salary information, personal phone numbers, and email addresses. Anthem, the US's second-largest health insurer, experienced a data breach in which the data of over 80 million former and current Anthem affiliates were stolen (Koerner 2016). The data stolen did not include private health records or credit card numbers, but rather seemed to target personal identification data including SSNs, income data, birthdays, street and email addresses. In a similar attack on Community Health Systems in August 2014, the personal information of over 4.5 million patients' data was stolen (Community Health Systems 2014). Again, the breach did not target intellectual property or financial or medical information, but focused rather on the names, addresses, birth dates, telephone numbers and SSNs of clients were stolen. And in May 2015, it was discovered that Chinese state-backed actors, again likely the PLA, had been accessing the United Airlines' database since April 2014 (Khandelwal 2015). The breach compromised information concerning flights, passengers and their movements, including passenger names, date of birth, departure and arrival locations. The breach of the Marriott Starwood hotel chains' database, discovered in September 2018, found that Chinese state-backed actors had been accessing the database since 2014, compromising up to 383 million travel records. The records included full names, phone numbers, email addresses, and date of birth, in addition to the 5 million passport numbers that were also exposed (Human Rights Watch 2017).

Data breach	Type of data	Potential uses
OPM: Attack 1: discovered March 2014. Attack 2: May 2014; Discovered April 2015	SF-86 background information on up to 4.1 million former/current employees, including: full names, job history, relationships, personal finances, past substance abuse/ psychiatric care, etc.; Fingerprints of 5.6 million government employees; 21.5 million SSN numbers	Counter-intelligence: rich in detail about persons of interest—previous workplaces, names of colleagues, foreign contacts, where they travel; Potential for blackmail.
Marriott Starwood: 2014; Discovered September 2018	Travel information on up to 383 million records lost, including: full names, phone numbers, email addresses, and date of birth; 5 million passport numbers.	Names can be matched with information from OPM. With passport data and birth names, the travel history of an individual could be pieced together.
Anthem Insurance: April 2014; discovered Jan 2015	Data on 80 million employees and members of Anthem, *not* involving private health records or credit card numbers, but exposing SSNs, income data, birthdays, and street and email addresses.	Bolsters list of information from OPM hack.

Navy Personnel: 2006 to 2018; Indicted in Dec 2018	Information on over 100,000 Navy personnel, including: the names, SSNs, dates of birth, salary information, personal phone numbers, and email addresses	Personal info of military personnel
Community Health Systems: August 2014	4.5 million clients' data stolen, including: names, addresses, birth dates, telephone numbers and Social Security numbers	Names and personal info
United Airlines: April 2014 Discovered May 2015	Data concerning flights' passengers and their movements, including names, their date of birth, and their departure and arrival locations.	Flight data could be used to cross-reference travel patterns of persons of interest

*Figure 1: A catalog of data breaches thought to be perpetrated by Chinese state-backed actors (*DHS Flight Tracker).

Several characteristics unique to these cyber-attacks point to the direct involvement of the Chinese government. Evidence from several data breaches implicate the Chinese military establishment, the People's Liberation Army. In an investigative report by Cybersecurity firm Mandiant, researchers found conclusive evidence that implicated PLA Unit 61398, the mission focus of which is signals intelligence, foreign language proficiency, and defense information systems (Mandiant). The report came to this conclusion for the following reasons: the IP addresses of several of the data breaches were traced back to China, specifically a PLA-operated building in Shanghai which had special fiber optic communications infrastructure installed for "national security" purposes (Mandiant). In addition, 97% of the 1,905 intruders observed by Mandiant in their post-breach investigation had their IP addresses registered in Shanghai, with language keyboards set to use Simplified Chinese. The ability to conduct such a long-running and extensive cyber espionage campaign also suggests state-backed support.

Evidence from the post-breach investigations of OPM and Anthem Health Insurance further implicate PLA Group 61398. In the OPM and Anthem hacks, PlugX was used; this was the same backdoor tool that had previously been used by the Chinese hacking group to target political activists in Hong Kong and Tibet (Koerner 2016). Similarly, the investigation into the Marriott Starwood breach revealed that the hacking tools, techniques, and procedures were the signature of the same group (Bing 2018). And as mentioned previously, the domain of "opmsecurity.org" was registered to "Steve Rogers," member of the Marvel Comic the *Avengers,* and a signature of Unit 61398.

Some, including the Communist Party of China (CCP), have claimed that the cyber-attack could be attributed to a cyber-criminal organization outside of the central government's control (Carsten 2015). While it is conceivable that a cybercriminal organization would be motivated to steal personal information including SSNs and personal contact information, three characteristics further point to the command of government resources.

1. The attackers maintained access to these databases for extended periods—many months, and years in some incidences. The long and sustained nature of the data breaches indicate substantial resources available to the attackers.

2. None of the data has been published. Criminal actors, motivated by financial gain, would be motivated to sell the personal information obtained from these data breaches. At the time of this writing five years have passed from the discovery of the initial data breaches; however, none of this data has surfaced on the dark web or used for financially motivated crimes.

3. The type of information stolen from the database indicate the support of a nation state. Although the databases targeted included financially valuable information, the data breaches did not include information like credit card numbers, which would be the target of any financially motivated cyber-criminal. Considering the personal information that could have been gained from the Marriott Starwood or the Anthem Insurance data breaches, the selection of data that was compromised indicates a lack of financial motivation.

Given the evidence and characteristics of these data breaches, the PLA intends to use this data to form a database on American citizens. This report makes the case that this database will be utilized by the Ministry of State Security (MSS) through continued data collection operations and advances in big data processing. The database will be a critical resource in the MSS's efforts to identify, target, and recruit American citizens to serve as informants for commercial, military, and political intelligence.

II. Goals of the threat actor

Understanding the motivations and intentions of the threat actor is critical in order to identify vulnerable data types of data that the threat actor may target next, and to devise countermeasures to prevent and mitigate the effects of the threat actor's data collection efforts. By identifying the Chinese Communist Party (CCP)'s overarching foreign policy and domestic goals, in addition to understanding operations of the Ministry of State Security and the People's Liberation Army, the we may begin to understand which types of data may be targeted in the future, and what can be done to mitigate this threat.

The Chinese Communist Party has evolved significantly since the establishment of the People's Republic of China (PRC), as have its foreign policy and domestic goals. China's long-term goals are shaped by its history, of which its "century of humiliation" plays a critical role. Following thousands of years of dynastic rule, the century of humiliation began in the mid 19th century with the Opium Wars, lasting until 1949, the founding of the People's Republic of China. This era was marked by continuous foreign occupation by Western colonial powers and Japan, with several failed attempts to reinstate Chinese control. In 1949, the CCP gained control over China and established the People's Republic of China (PRC). The legacy of the century of humiliation is evident from its two "hundred year goals", which set to (1) build a moderately prosperous society by 2021, when the CCP celebrates its centenary, and (2) build a modern socialist country that is prosperous, strong, democratic, culturally advanced and harmonious by 2049. China's development goals are also reflected in President Xi Jinping's "China Dream", which aspires for the "great rejuvenation of the Chinese nation" (Pillsbury 2016). Despite assurances from President Xi that China "will not seek to dominate", China's plans for revitalization may put the country at odds with the U.S (Xi Jinping 2018).

The fundamental goal of the CCP is to maintain control and domestic stability. China's leaders seek to expand China's growing economic, diplomatic and military presence in an effort to establish regional preeminence, and to expand the country's influence internationally (The State Council 2015). As geopolitical tensions in the Asia-Pacific intensify, China's military has stated its determination in safeguarding its interests in this region, and to safeguard and counter the US's "rebalancing" strategy in the region (The State Council 2015). Another central goal of the CCP is to maintain its legitimacy. The CCP secures its legitimacy from its ability to provide economic growth and stability within China; without it, it fears that instability and threats to the central party will follow. This concern, coupled with the insecurity stemming from the century of humiliation, explain China's motivations as a revisionist power. According to President Xi Jinping's long-term plans, China should be a top-ranked nation in innovation by 2035, and by 2050, China should become a nation with pioneering global influence. China's 13th five-year plan (2016-2020) calls for greater technology innovations and socioeconomic reform. The "Made in China 2025" plan, the AI Development Strategy are just two more of several initiatives to expand China's global influence and rise as an economic leader. China is also striving to expand its soft power, evident in its development projects throughout sub-Saharan Africa and South Asia (Heath 2018).

China's overarching goals shape the motivations and usage for the data collected. There is obvious

intelligence value for any country—ally or enemy—to be gathered from the hacks on OPM, Marriott Starwood, United Airlines, and various healthcare entities. But in the eyes of the CCP, the United States is a threat to China's goals for regional dominance and expansion internationally. From US Naval patrols of disputed territories in Pacific waters, to the US's continued support of Taiwan, to US efforts to obstruct the business of Chinese companies like Huawei and ZTE, the US has indicated that it may present barriers to China's goals. As the CCP is set on rejuvenation and increasing its standing in the world, the current global hegemon, by its actions and statements, may be perceived as a barrier to achieving the "China Dream". One incidence of this might include the ban of export of computer chips from the US to China, thereby obstructing the development and production cycles of supercomputing companies. It would be feasible for the Chinese to seek to obtain other methods to obtain access to the knowledge and production of these computer chips. The collection of data on US citizens enables the Chinese state to identify and categorize US citizens of interest, whereupon they may be targeted for intelligence gathering purposes.

III. Operations, capabilities, and modus operandi of China's state intelligence agencies

i. PLA Unit 61398 and the Strategic Support Force

The cyber-attacks targeting US citizens' data has been attributed to the People's Liberation Army. Specifically, the PLA Unit 61398 has been identified as the culprit in several of these attacks, including OPM and Anthem Health Insurance. Unit 61398, whose official name is China's Military Unit Cover Designator (MUCD) 61398, functions as the PLA's cyber command. As shown in Figure 2, the PLA reports directly to the CPC's Central Military Commission. The PLA's cyber command, including Unit 61398, fall under the PLA's 3rd General Staff Department's 2nd bureau. The 3rd General Staff department's focus is on signals intelligence, foreign language proficiency, and defense information systems; it is likely that those working in the 2nd bureau have been responsible for the attacks focusing on gathering data on US citizens, but other goals include economic and industrial espionage (Mandiant). Publicly available resources confirm that Unit 61398's mission focus is on computer network operations, and a report on China's signals intelligence infrastructure from the Project 2049 Institute found that Unit 61398's targets included the US and Canada, with a focus on "political, economic, and military-related intelligence" (Stokes et al. 2011).

For a period of 18 months from 2015 to 2017, cyber offensive groups from China seemed to become less active. In June 2016, FireEye reported dramatic decreases in activity from 72 suspected China-based cyber espionage groups since 2014 (Mandiant). Reasons for this could include a bilateral agreement reached between Presidents Barack Obama and Xi Jinping on cyber espionage in September 2015. While the two governments agreed that they would not conduct or knowingly support cyber-enabled commercial IP theft, the two countries did not agree to cease government espionage, which is a generally accepted activity.

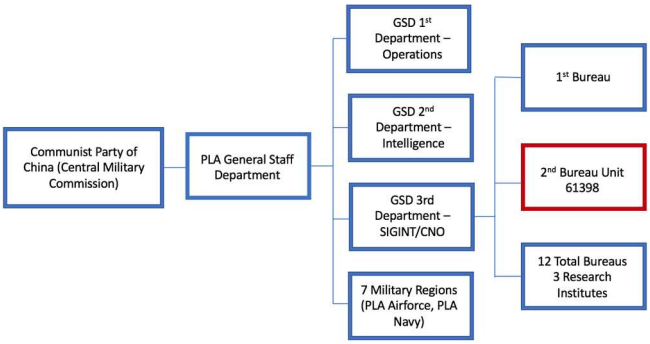

Figure 2: Organizational Structure of the People's Liberation Army's GSD 3rd Department. Figure adapted from Mandiant. Feb 2013. "APT 1: Exposing One of China's Cyber Espionage Units."

For a period of 18 months from 2015 to 2017, cyber offensive groups from China seemed to become less active. In June 2016, FireEye reported dramatic decreases in activity from 72 suspected China-based cyber espionage groups since 2014 (Mandiant). Reasons for this could include a bilateral agreement reached between Presidents Barack Obama and Xi Jinping on cyber espionage in September 2015. While the two governments agreed that they would not conduct or knowingly support cyber-enabled commercial IP theft, the two countries did not agree to cease government espionage, which is a generally accepted activity.

The PLA's ambitious military modernization and organizational reforms may better explain the

decrease in activity. The establishment of the Strategic Support Force (SSF) reflects an innovation in the military's force structure, which ultimately seeks to optimize China's capabilities in the space, cyberspace and electromagnetic domains (Office of the Secretary of Defense 2018). The centralization of the cyber warfare command under the Central Military Commission (CMC) can be seen as part of a greater effort to consolidate and optimize the PLA's capabilities in order to fight and win future "informatized" wars (Kania et al. 2018).

Following the brief hiatus of cyber offensive operations, several cybersecurity firms including CrowdStrike and FireEye reported a resurgence in cyberespionage efforts stemming from the PLA (Johnson 2018; FireEye 2016). However, the intrusions have become more difficult to detect. Increasingly common is the use of generic "tools", leaving limited to no unique signatures, making attribution difficult (Johnson 2018). While some of this resurgence could be attributed to worsening trade relations between the US and China, the lull and increase in discreet cyber intrusions is likely to be a result of the PLA's cyber force restructure. The establishment of the SSF indicates shifts in the threat actor's operations, which will be examined in the next section of this report.

ii. Coordination within China's intelligence agencies

While the PLA/SSF conduct the bulk of the cyber offensive operations to collect data for the database, the Ministry of State Security (MSS) is responsible for interpreting and utilizing the information collected. As China's main foreign intelligence service, the MSS's efforts in gathering human intelligence is one of its key objectives. China's new National Intelligence Law, passed in June 2017, sought to facilitate cooperation between state intelligence agencies by establishing a "state intelligence work coordination mechanism" (Hoffman et al. 2017). This cooperation may also be enabled due to the central command structure of intelligence organizations within China (see Figure 3). The coordination between the PLA and MSS may further be supported by Xi Jinping's drive for greater integration between the intelligence services within China and in the CCP's push to integrate its cyberwarfare capabilities.

The Ministry of State Security follows some key operating procedures. First, the MSS conducts most of its intelligence operations from within mainland China. The limited intelligence networks the Ministry of State Security has abroad is rooted in the 1970's when President Deng Xiaoping banned the use of cover posts in diplomatic missions for being used for espionage purposes (Eftimiades 1994). This is reflected in Taiwan, where China has been most successful in establishing an informant network. Of more than a dozen Chinese espionage cases that were studied from 2010-2014, only one occurred outside of China, and this case remained an anomaly regarding the MSS's normal operating procedure (Mattis 2014).

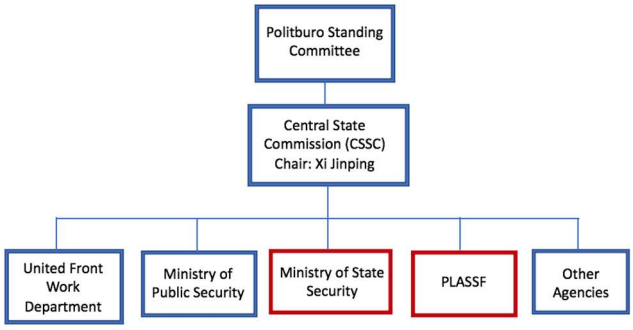

Figure 3: Command structure of Chinese intelligence agencies under the CSSC. Figure adapted from Jane's By HIS Markit. 2017. "Chinese Legislation Points to New Intelligence Coordinating System."

The stories of two American spy recruits further supports the theory that the MSS prefers to lure potential informants to mainland China before making an official proposal. Kevin Mallory, a former CIA official, was sentenced by a US federal court to 20 years in prison for attempting to provide classified documents to an agent of the PRC (US Department of Justice). Mallory was struggling financially when he was contacted via LinkedIn by a Chinese "headhunter". The MSS operative then arranged a phone call between Mallory and another individual, under the guise of a job with a think tank in Shanghai. Following two trips to Shanghai, Mallory agreed to sell defense secrets to his new Chinese contacts. American college student Glenn Shriver was likewise also recruited to spy for the MSS while in mainland China. During his study abroad in China, he responded to a newspaper ad asking for someone to write a white paper about trade relations between the US, North Korea, and Taiwan. He was approached by a woman who

offered him $120 for the essay and was subsequently recruited to become an informant for the MSS (Mattis 2015). Similar efforts to recruit Western nationals through social media sites like LinkedIn have also been reported by intelligence agencies in the United Kingdom and in Germany (Federal Ministry of the Interior 2016; Burgess 2015). These separate incidences further support the idea that the MSS operates mostly within its own territory. And while it is possible that recruitment occurs outside of mainland China, historical patterns of informant recruitment indicate that domestic outreach is critical to the operation of Chinese intelligence agencies.

Dozens of incidences of informant recruitment by Chinese intelligence services give credence to the theory that the MSS follows a step-by-step recruitment process (Graff 2018, Stratfor Worldview 2019; Aatola 2019).

Step 1: Spotting
Intelligence officials identify people of interest. The OPM database provides a wealth of data for this; the 4.1 million SF-86 background check files of former and current federal government employees include full names, full job histories, SSNs, and fingerprints. This gives intelligence services an idea of which people may be of interest for targeting. Combined with the Navy Personnel and Anthem Insurance databases, Chinese intelligence services can form a broad database of the type of careers select individuals have, and the type of information these individuals may have access to, in both the private and public sectors

Step 2: Assessing
Once intelligence officials identify potential recruits, they examine how those targets might be encouraged to spy. Common motivators include money, belief in the cause, blackmail, and ego. The SF-86 background check information from the OPM hack includes details on personal relationships, personal finances, past substance abuse, gambling addictions, psychiatric care etc. This type of information provides a comprehensive playbook with which to lure or coerce potential spy recruits.

Step 3: Developing
Having identified and assessed their potential targets, Chinese intelligence officials may then begin to groom their source to establish rapport. Having established that the Ministry of State Security operatives rarely recruit outside of mainland China, it is likely that this process will take place through virtual communications.

Step 4: Recruiting
Step 4 of the spy recruitment process, "recruitment" likely happens when the informant travels to mainland China. The informant may travel on their own or may be lured by Chinese intelligence services under the guise of a career-related reason. The data collected from the United Airlines and Marriott Starwood hacks could provide flight histories and passport information of several million Americans, informing the MSS on the travel patterns of these individuals

Step 5: Handling
Step 5 is the maintenance of the relationship between informant and the intelligence apparatus. This is the method with which the informant would relay information back to the Chinese intelligence agency, in addition to how the informant would receive further instructions. This step of the recruitment process maintains an already established link. Therefore, it is less likely to benefit from additional data/information on informants. Focusing on steps 1-4 of the recruitment process – identifying, targeting and recruitment of potential American informants, the report identifies several vulnerabilities that may be targeted as a part of China's informant recruitment process.

The evidence found in this report indicates the PLA and MSS will target data that will aid them in their informant recruitment process—who to target, how to target them, and when. The next section of the report will be dedicated to identifying data sources that China will target to bolster its database on American citizens, followed by recommendations to protect against these vulnerabilities.

IV. High-value strategic targets and policy recommendations

This report identifies three key data targets that would provide great utility for Chinese intelligence agencies for their informant recruitment process in the United States. These potential vulnerabilities can be mitigated through a set of solutions through the

use of counterintelligence measures, cybersecurity best practices, and data privacy standards.

i. Target: Data broker companies

Regarding Step 1 of the recruitment process, identifying potential informants, data that can provide basic knowledge on American individuals' names, jobs, and perhaps contact information would be most valuable. Armed with data from the OPM data breach, it is likely that Chinese intelligence sources would like to target other background information resources, including the CIA and NSA's employee databases in its search. However, Chinese intelligence agencies have indicated that their espionage goals are not limited to government agencies but includes the private sector in their targets as well. Data broker companies, particularly those that specialize in "people searches", hold a high value targets. These types of data brokers, including companies like Acxiom, Datalogix and PeekYou, gather third party information on millions of Americans including names, phone numbers, locations, emails, SSN, education information, job information, marital status and social media, providing a base from which Chinese could identify potential targets. The limited regulations regarding the types and amount of consumer information make it relatively easy for data broker companies to amass information. However, the cybersecurity standards which these companies are held to are relatively lax. This is evidenced by an analysis of the top 100 data brokers in the US, of which only 25% encrypted their landing pages, and 50% encrypted login pages. Most data broker companies only subscribe to security as a service offering, and security seals on their own are not effective countermeasures (Haynes 2017).

In a report on data brokers' collection of consumer information, the Federal Trade Commission found several practices regarding storage and retention of data that may impact the privacy and security of the consumers about whom the data was collected (Federal Trade Commission 2014). The FTC has also taken action against data broker companies Reed Elsevier and Seisint for security flaws that allowed identity thieves to exploit the companies' databases (Federal Trade Commission 2008). These cybersecurity vulnerabilities, coupled with the unrestricted collection of personal data, presents the perfect opportunity for PLA/SSF operatives to hack into the data brokers' databases without detection, contributing to the database on American citizens.

Policy recommendation 1: Implement stricter cybersecurity standards and privacy regulations for data broker companies

Currently, data broker companies' data collection methods are wholly legal, as they collect information from publicly available resources (including public records, commercial purchase history, and social media). Though broker companies are subject to the Fair Credit Reporting Act (FCRA), several lawsuits involving data broker companies indicate these companies are rarely held accountable for reporting incorrect information (Federal Trade Commission 2014). This highlights the ease with which data brokers may collect and disseminate information. The FTC recommended that Congress consider legislation to account for the privacy and security vulnerabilities, including the deletion of older data, and to allow consumer to opt out and possess greater propriety over the data that is stored on them, such as allowing consumers access to their data, and greater regulations on the collection and storage of sensitive data (DHS Flight Tracker). This practice is supported by the National Institute of Standards and Technology (NIST)'s Cybersecurity Framework, which posits that an organization's cybersecurity activities creates risks when personal information is collected and used without consideration for privacy, and further the over-collection or over-retention of personal information may result in heightened security risks (NIST 2018). Several states have begun to regulate the data collection and cybersecurity standards of data broker companies. From January 2019, the state of Vermont required that data brokers adopt comprehensive security measures, and to publicly disclose the companies' data collection practices, opt-out policies purchaser credentialing practices, and security breaches (Goldstick et al. 2019).

This report suggests that the US federal government establish a regulatory law concerning the data collection practices and cybersecurity standards of data broker companies. By addressing both data privacy and cybersecurity vulnerabilities, the risk of an intrusive data breach could be reduced.

ii. Target: Open-source social media (LinkedIn)

Regarding steps 2 and 3 of the recruitment process, the "assessment" and "development" of a potential spy recruit, Chinese intelligence operatives will seek to gain a more in-depth understanding of potential spy recruits, and subsequently attempt to contact these individuals. Open-source social media, such as LinkedIn, would provide valuable background information on individuals, in addition to providing a medium with which to "develop", or "groom" individuals for espionage. Social media sites, such as LinkedIn, would be ideal for the MSS to reach out and communicate with prospective informants. The German intelligence service reported that MSS operatives, posing as headhunters, targeted over 10,000 German politicians, scientists, and other professionals through LinkedIn. These headhunters reach out to people over LinkedIn, after which they "luring [them] with enticing offers and eventually inviting [them] to China, where the intelligence-gathering commences" (Federal Ministry of the Interior 2016).The German intelligence service reported that Chinese espionage efforts focused on industry, research, technology and the armed forces, in addition to gathering intelligence on German political processes, specifically anything that may pose a threat to the CCP's monopoly on power (Federal Interior of the Minister 2016). In the UK, the MI-5 released a memo warning government workers that Chinese operatives were utilizing LinkedIn social network to target government employee (Burgess 2015). William Evanina, the US counter-intelligence chief, also confirmed that Chinese intelligence agencies were also using fake LinkedIn accounts to recruit Americans with access to government and commercial secrets (Strobel et al. 2018). The fact that Chinese intelligence officers rarely operate outside of mainland China increases the likelihood that they would take advantage of openly available social media connections.

Policy recommendation 2: Counterintelligence and user-based cybersecurity recommendations for social media platforms

LinkedIn is arguably one of the most useful social media platforms with which Chinese intelligence services may use to search, contact and recruit potential informants and other persons of interest. Knowing that (1) China's intelligence agents work almost exclusively within the geographic confines of mainland China, and (2) several past incidences of informant recruitment have begun online, the solution recommends counterintelligence measures be taken against these so-called headhunters, in addition to recommending user-based security practices.

First, identify and publicize fake profiles of Chinese "headhunters". US counterintelligence services should work to identify and publicize the accounts of the fake headhunters operating on LinkedIn. The Clarifying Lawful Overseas Use of Data Act, or CLOUD Act, is a federal law enacted in 2018, which established processes and procedures for US cloud service providers to comply to law enforcement requests for access to data in other countries, if a warrant or subpoena exists (Kris 2015). The CLOUD Act, supported by the Department of Justice as well as by major tech companies (including Microsoft, Apple, and Google), would legally compel LinkedIn to aid the US government in identifying these users. LinkedIn has already complied to government requests to remove fake profiles in the past, including the deactivation of LinkedIn accounts that German officials had identified as spies (Hernandez et al. 2017). LinkedIn may be able to pinpoint the actual identities of the headhunters on LinkedIn from the "verification" data it gathers on users in China, courtesy of China's Cybersecurity Law, passed in 2017 (Liao 2019).

This law requires users to verify their identities through their phone numbers and a "real-name verification process." These legal requirements, part of an effort to end digital anonymity in China, may give LinkedIn hints of the true identities of the headhunters. While these users are likely to provide fraudulent identities, the phone numbers could be traced to some organization or individual. With cooperation between the US government and LinkedIn, the threat posed by China's intelligence sources to communicate with and recruit individuals may be mitigated.

Second, develop user-based best practices. LinkedIn users should also be warned about the suspicious behavior of fake headhunters operating on LinkedIn, as well as the potential consequences of consorting with these Chinese informant recruiters. Warning social media users about this threat may fall in line with similar efforts to educate social media users how to other fake accounts, such as with Russian

bots and trolls during the US Presidential Election in 2016 (Aneia et al. 2018).

By taking an active stance on identifying, removing and warning about the threat of Chinese operatives operating on LinkedIn, the vulnerability of openly available data resources may be reduced. Effectively enforcing these measures will make it more difficult for China's intelligence officials to assess, target and communicate with potential informants.

iii. Vulnerability: DHS Flight Tracker

Step 4 of the spy recruitment process, the actual "recruitment", generally occurs when individuals travel to China. Through the United Airlines and Marriott Starwood hacks, Chinese intelligence services likely have information on the flight patterns of United passengers, in addition to five million passport numbers. This information will be key in the MSS knowing when and where to engage with persons of interest. The DHS's flight tracker contains passport numbers and the arrival and departure flight history for individuals' dating five years back. Combined with the passport information stolen from the Marriott Starwood hack, Chinese intelligence officials could cross reference the information from the DHS flight tracker to identify the travel patterns of targeted individuals (DHS Flight Tracker).

Policy recommendation 3: Strengthening cyber defenses of the DHS Flight Tracker

The DHS should be advised that its information has high strategic value for Chinese intelligence agencies. The department should seek to bolster its cyber defense infrastructure, in addition to increase its efforts to detect malicious intruders in the database. With the SSF's operations becoming increasingly well-concealed, it is likely that an intrusion may not be noticed until it is too late. And while a diplomatic solution, like that of the 2015 agreement to cease commercial IP theft resulted in a slowdown in hacking activity, it is unlikely that either country would come to a agreement to halt intelligence gathering operations. Coupled with inherent difficulties with attribution and the current state of US-China relations, a cyber cease-fire in this situation is unlikely.

V. Challenges and limitations to analysis and implementation of these recommendations

Due to the nature of cyber espionage and the analysis of intelligence issues, it is impossible to fully understand or be certain of the threat actor's goals or plans, nor is it possible to devise detailed security plans for the suggested targets in the previous section. While publicly available statements from the PLA and from independent analysts suggest that coordination efforts between the PLA and MSS is likely, there has been no official confirmation from the Chinese government or from US intelligence services. In addition, the cybersecurity vulnerabilities of the DHS are not disclosed to the public for obvious reasons, but as a result the report cannot anything more than a blanket advisory for the DHS to defend against this threat. Further, the views portrayed in this report are based off a review of publicly available government publications and statements, journal articles, and news articles. Although the evidence presented in this report suggested certain behaviors and motivations, this report by no means claims all-encompassing understanding of the motivations and operations of the Chinese government. Security experts should consider the threat posed to forms of media that could provide personal information about individuals, particularly data that could be used to blackmail individuals (Reddit accounts, dating app information, etc.).

Additionally, the recommendations presented in this report may encounter legal challenges. The Federal Trade Commission recommended greater transparency, accountability, and cybersecurity standards for data broker companies five years ago. Despite progress on this issue in the state of Vermont, there will likely be challenges to enacting federal regulations on this issue. And although LinkedIn has shown a willingness to remove and publicize the identities of the MSS operatives recruiting Western intelligence sources, China's Cybersecurity Law and the Personal Information Security Specification (2018) requires firms to store data locally in China, thereby preventing some of the information regarding flagged profiles to be shared with Western governments (Kirkpatrick 2018).

VI. Conclusion

While none of the data stolen in the OPM, Marriott Starwood, and United Airlines hacks have had

immediately damaging effects to individuals, this report found that this data may be used to identify, target, and recruit US citizens as spies for the Chinese state. This use case would present a significant threat to the national security of the United States and its citizens. Amidst increasing tension between the US and China and the revamping of the Strategic Support Force, it is likely that we will see more utilization of this database in the future. The US government and its citizens need to understand the magnitude of this threat. While this report identifies key strategic targets and recommendations, there are still countless other databases that can, and will, be targeted. With an understanding of what Chinese intelligence operations are hoping to accomplish, the US can be better equipped to mitigate this threat.

References

Aaltola, Mike. 2019. "Geostrategically Motivated Co-option of Social Media." Finnish Institute of International Affairs. https://www.fiia.fi/wp-content/uploads/2019/06/bp267_geostrategically_motivated_co-option_of_social-media.pdf

Aneja, Arpita, Sandra Ifraimova,. 2018. "How to Spot a Russian Bot." Time Magazine. http://time.com/5274785/how-to-spot-a-russian-troll/

Armerding. 2016. "The OPM Breach Report: A Long Time Coming." CSO Online. https://www.csoonline.com/article/3130682/the-opm-breach-report-a-long-time-coming.html

BBC News. 2018. "Xi Jinping says China 'will not seek to dominate.'" BBC News. https://www.bbc.com/news/world-asia-china-46601175

Bing, Christopher. "Exclusive: Clues in Marriott Hack Implicate China – Sources." Reuters. https://uk.reuters.com/article/uk-marriott-intnl-cyber-china/clues-in-marriott-hack-implicate-china-sources-idUKKBN1O504B

Burgess, Christopher. 2015. "Beware where you share: British Intelligence Cautions Employees Against LinkedIn." Clearance Jobs. https://news.clearancejobs.com/2015/08/21/beware-share-british-intelligence-cautions-employees-linkedin/

Carsten, Paul, Mark Hosenball. 2015. "China's Xinhua says US OPM hack was not state-sponsored." Reuters. https://www.reuters.com/article/us-china-usa-cybersecurity/chinas-xinhua-says-u-s-opm-hack-was-not-state-sponsored-idUSKBN0TL0F120151202

Community Health Systems, Inc. 2014. "United States Securities and Exchange Commission." Sec.gov. https://www.sec.gov/Archives/edgar/data/1108109/000119312514312504/d776541d8k.htm

DHS Flight Tracker. https://i94.cbp.dhs.gov/I94/#/home

Nicholas Eftimiades. 1994. Chinese Intelligence Operations. Naval Institute Press.

Federal Ministry of the Interior. 2016. "Brief Summary: 2016 Report on the Protection of the Constitution." https://www.verfassungsschutz.de/embed/annual-report-2016-summary.pdf

Federal Trade Commission. 2008. "Agency Announces Settlement of Separate Actions Against

Retailer TJX, and Data Brokers Reed Elsevier and Seisint for Failing to Provide Adequate Security for Consumers Data." Federal Trade Commission. https://www.ftc.gov/news-events/press-releases/2008/03/agency-announces-settlement-separate-actions-against-retailer-tjx

Federal Trade Commission. 2014. "FTC Recommends Congress Require the Data Broker Industry to be More Transparent and give Consumers Greater Control over their Personal Information." FTC. https://www.ftc.gov/news-events/press-releases/2014/05/ftc-recommends-congress-require-data-broker-industry-be-more?utm_source=govdelivery

Ambassador Chas W. Freeman. "China's Challenge to American Hegemony: Remarks to the

Global Strategy Forum." Middle East Policy Council. https://www.mepc.org/speeches/chinas-challenge-american-hegemony

FireEye. 2016. "Redline Drawn: China Recalculates its Use of Cyber Espionage." Fireeye ISight Intelligence. https://www.fireeye.com/content/dam/fireeye-www/current-threats/pdfs/rpt-china-espionage.pdf

Goldstick, Samuel D, Jennifer L Rathburn, Aaron K Tantleff. "Ringing in 2019 with New State Privacy and Data Security Laws Impacting Data Brokers and Insurers." Foley and Lardner LLP. http://www.mondaq.com/unitedstates/x/771870/Security/Ringing+in+2019+with+New+State+Privacy+and+Data+Security+Laws+Impacting+Data+Brokers+and+Insurers

Graff, Garrett M. 2018. "China's 5 Steps for Recruiting Spies." Wired. https://www.wired.com/story/china-spy-recruitment-us/

Haynes, Alex. 2017. "Are Data Brokers Actually Secure?" Info Security. https://www.infosecurity-magazine.com/opinions/are-data-brokers-actually-secure/

Heath, Timothy. 2018. "China's Pursuit of Overseas Security." RAND Corporation. https://www.rand.org/content/dam/rand/pubs/research_reports/RR2200/RR2271/RAND_RR2271.pdf

Hernandez, Javier C and Melissa Eddy. 2017. "Germany Accuses China of Using LinkedIn to Recruit Informants." The New York Times. https://www.nytimes.com/2017/12/11/world/asia/china-germany-linkedin.html

Hoffman, Samantha and Peter Mattis. 2017. "Chinese Legislation Points to New Intelligence Coordinating System." Jane's By IHS Markit. https://www.janes.com/images/assets/183/74183/Chinese_legislation_points_to_new_intelligence_co-ordinating_system.pdf

Human Rights Watch. 2017. "China: Police 'Big Data' Systems Violate Privacy, Target Dissent." Human Rights Watch. https://www.hrw.org/news/2017/11/19/china-police-big-data-systems-violate-privacy-target-dissent

Xi Jinping's speech marking the 40th anniversary of the country's reform and opening up to the market economy, at the Great Hall of the People on December 17, 2018. https://www.youtube.com/watch?v=MILBtNHX4rQ

Johnson, Tim. 2018. "China Backed Off From Hacking US Companies. Now it is at it again." McClatchy. https://www.mcclatchydc.com/news/nation-world/national/national-security/article212666139.html

Kania, Elsa B, John K Costello. 2018. "The Strategic Support Force and the Future of Chinese Information Operations." The Cyber Defense Review. https://cyberdefensereview.army.mil/Portals/6/Documents/CDR%20Journal%20Articles/The%20Strategic%20Support%20Force_Kania_Costello.pdf?ver=2018-07-31-093713-580

Khandelwal, Swati. 2015. "United Airlines Hacked by Sophisticated Hacking Group." The Hacker News. https://thehackernews.com/2015/07/united-airlines-hacked.html

Kirkpatrick, Keith. 2018. "Borders in the Cloud." Communications of the ACM https://cacm.acm.org/magazines/2018/9/230563-borders-in-the-cloud/fulltext

Koerner, Brendan L. 2016. "Inside the Cyberattack that Shocked the US Government." Wired. https://www.wired.com/2016/10/inside-cyberattack-shocked-us-government/

Kris, David. 2015. "Preliminary Thoughts on Cross-Border Data Requests." Lawfare. https://www.lawfareblog.com/preliminary-thoughts-cross-border-data-requests

Liao, Rita. 2019. "LinkedIn Now Requires Phone Number Verification for All Users in China." Tech Crunch. https://techcrunch.com/2019/01/09/linkedin-real-name-phone-number-verification-china/

Mandiant. "APT 1: Exposing One of China's Cyber Espionage Units." Mandiant. https://www.fireeye.com/content/dam/fireeye-www/services/pdfs/mandiant-apt1-report.pdf

Mattis, Peter. "China's Espionage Against Taiwan (Part 1: Analysis of Recent Operations)" The Jamestown Foundation. https://jamestown.org/program/chinas-espionage-against-taiwan-part-i-analysis-of-recent-operations/

Mirani, Leo, Max Nisen. 2014. "The nine companies that know more about you than Google or Facebook." Quartz. https://qz.com/213900/the-nine-companies-that-know-more-about-you-than-google-or-facebook/

Nakashima, Ellen, David J Lynch. 2018. "U.S. charges Chinese Hackers in Alleged Theft of Vast Trove of Confidential Data from 12 Countries." The Washington Post. https://www.washingtonpost.com/world/national-security/us-and-more-than-a-dozen-allies-to-condemn-china-for-economic-espionage/2018/12/20/cdfd0338-0455-11e9-b5df-5d3874f1ac36_story.html

National Institute of Standards and Technology. 2018. "Framework for Improving Critical Infrastructure Cybersecurity, Version 1.1." https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.04162018.pdf

Office of the Secretary of Defense. 2018. "Annual Report to Congress: Military and Security Developments Involving the People's Republic of China 2018." Department of Defense. https://media.defense.gov/2018/Aug/16/2001955282/-1/-1/1/2018-CHINA-MILITARY-POWER-REPORT.PDF

Michael Pillsbury 2016. The Hundred Year Marathon: China's Secret Strategy to Replace American as the Global Superpower. St. Martin's Press.

The State Council Information Office of the People's Republic of China. 2015. "China's Military Strategy." http://eng.mod.gov.cn/Press/2015-05/26/content_4586805.htm

Stokes Mark A, Jenny Lin, L.C. Russell Hsiao. 2011. "The Chinese People's Liberation Army

Signals Intelligence and Cyber Reconnaissance Infrastructure," Project 2049 Institute (2011): 8 http://project2049.net/documents/pla_third_department_sigint_cyber_stokes_lin_hsiao.pdf

Stratfor Worldview. 2019. "Beware: Iran and China Use LinkedIn to Recruit Spies." The National Interest.

https://nationalinterest.org/blog/buzz/beware-iran-and-china-use-linkedin-recruit-spies-65761

Strobel, Warren and Jonathan Landay. 2018. "Exclusive: US accuses China of 'super aggressive' spy campaign on LinkedIn." Reuters. https://www.reuters.com/article/us-linkedin-china-espionage-exclusive/exclusive-us-accuses-china-of-super-aggressive-spy-campaign-on-linkedin-idUSKCN1LG15Y

Threat Connect Research Team. 2015. "OPM Breach Analysis: Update." Threat Connect. https://threatconnect.com/blog/opm-breach-analysis-update/

United State Department of Justice. 2019. "Former CIA Officer Sentenced to Prison for Espionage." US Department of Justice. https://www.justice.gov/opa/pr/former-cia-officer-sentenced-prison-espionage

Ming (Sherry) Chen is a recent graduate of the Master's in Cybersecurity Policy program at Georgia Tech, where she also received her Bachelor's in International Affairs. Her interests include technology and policy, especially as they pertain to national security. Her previous work includes research on AI and technology developments in China and Russia.

Acknowledgments

The author would like to acknowledge Dr. Jaclyn Kerr and the staff at the Center for Global Security Research for their guidance and support, and Dr. Milton Mueller for his contributions in the preparation of this manuscript.