RICE UNIVERSITY

Analyzing the Use of Cyber in Warfare at the Strategic, Operational, and Tactical Levels

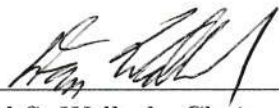
by

Judson Clark Dressler

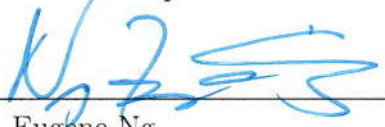
A Thesis Submitted
in Partial Fulfillment of the
Requirements for the Degree

Doctor of Philosophy

Approved, Thesis Committee:



Daniel S. Wallach, Chair
Professor of Computer Science



T. S. Eugene Ng
Associate Professor of Computer Science

Christopher Bronk
Assistant Professor of Computer Science

Houston, Texas

May, 2015

ABSTRACT

Analyzing the Use of Cyber in Warfare at the Strategic, Operational, and Tactical Levels

by

Judson Clark Dressler

The United States relies on networked computing for all manner of economic, social, and civic activity. However, cyberspace also presents potential adversaries with an avenue to overcome the overwhelming advantage enjoyed by the US in conventional military power. The introduction of cyberspace has blurred the edge of the battlefield; allowing an adversary to use easily procured equipment and from anywhere attack the process of a commercial or government target. This addition has introduced challenges to many traditional military concepts at each level of warfare: strategic, operational, and tactical.

This thesis investigates and presents solutions to three of these challenges.

At the strategic level, the DoD has declared cyberspace as a war-fighting domain. The ultra high-speed, fluid, and omnipresent nature of cyberspace makes it fundamentally different from the traditional domains. Strategic thinkers cling to ideological legacies of the past regarding problems, innovations, and strategies. So before imposing past tenets of and terminology onto the new field, these legacies need to be examined to see if they are pertinent and to what degree. This thesis debunks previous ideological molds as they pertain to cyberspace and ensures correct terminologies and frameworks are used, providing the DoD with a better understanding of how cyber possibly fits into the domain of warfare.

At the operational level, the DoD relies heavily on networking technologies to

efficiently conduct missions across the globe. This dependency places the nation at risk of a loss of confidentiality, integrity, and availability of its critical information resources, degrading its ability to complete the mission. I introduce the operational framework for establishing situational awareness in cyberspace. Using this framework will provide the nation's leadership timely and accurate information to gain an understanding of the operational cyber environment to enable decision-making at all levels. The DoD has already begun integrating this framework into an operational situational awareness tool.

At the tactical level, there has become a growing tension between military users' personal needs and military operational security in regards to use of social media. Like everyone, military members post seemingly trivial information and pictures, which can be aggregated and augmented by an adversary to determine possible intelligence targets. I investigate the current state of DoD social media policy, use an automated approach to determine the amount of openly available information provided by U.S. military members, analyze it through content analysis, apply machine learning techniques, then finally rank the vulnerability of each individual. In all, 1168 potential intelligence targets were discovered, of which, 223 were determined to be vulnerable. Most importantly, I demonstrated that automated methods can be effective in discovering easily targetable personnel by an adversary.

ACKNOWLEDGEMENTS

Though only my name appears on the cover of this dissertation, a great many people have contributed to its production. I owe my gratitude to all of those people who have made this dissertation possible.

My deepest gratitude to my advisor, Dr. Dan Wallach. I have been amazingly fortunate to have an advisor who gave me the freedom to explore on my own, switch gears when necessary, and work with me through all the difficulties that arrived. His patience and support helped me overcome a rocky start and a shortened time line to finish this dissertation. I will be forever grateful for this opportunity.

My co-advisor, Dr. Christopher Bronk, has always been there to listen and give advice. Under his guidance, I delved deep into the collective understanding of cyber policy past and present and how it may affect the field in the future, including my role in the military. His insightful comments and constructive criticisms at different stages of my research were thought-provoking and helped me remain calm and focused, while his numerous discussion tangents allowed me to stay sane.

Dr. T.S. Eugene Ng is one of the best teachers that I have had in my life. He sets high standards for his students and encourages and guides them to meet them. I am grateful to him for holding me and my research to such a high quality and for his continuous encouragement.

I would like to thank the United States Air Force, United States Air Force Academy, and in particular its Department of Computer Science. Colonel David Gibson and Dr. Martin Carlisle not only provided me the opportunity to obtain my PhD, but provided the foundation of my understanding of computer science and security. I look forward to returning back to the 'Zoo' and continuing to prepare the

future leaders of our Air Force.

I am also indebted to the members of the Security Group with whom I have interacted during the course of my graduate studies. Particularly, I would like to acknowledge Theodore Book, Adam Pridgen, and Dr. Michael Dietz for the many valuable discussions that helped me understand my research area better and in prepping me for the C-exam.

I would like to acknowledge Dr. Christopher Jermaine, Dr. Genevera Allen, and Dr. Erzsebet Merenyi for numerous discussion and lectures on related topics that helped me improve my knowledge in the area.

I would like to acknowledge all of my teachers throughout my life who have believed in me and pushed me to believe in myself, it was their backing that has propelled me to these heights.

To my friends and co-authors - especially Dr. Steve Fulton, David Merritt, Clay Moody, Triip Bowen, and Jason Koepke - thank you for guiding me and encouraging me in this endeavor.

I am also grateful to the following former and current staff at Rice University, for their various forms of support during my graduate study – Dr. Ron Goldman, Belia Martinez, Beth Rivera, Shery Nassar, and Lena Sifuentes. You all have been instrumental in working through the unique challenges of my presence at Rice and allowed me to stay focused on my research.

Most importantly, none of this would have been possible without the love and patience of my family. My parents – who dedicated their lives to improving mine, who taught me to work hard, to love God, and to always do the right thing; my children – who didn't always understand why I was couped up in my office or preoccupied thinking through thousands of lines of code; and finally, my wife Sarah – whose support, encouragement, quiet patience and unwavering love were undeniably the bedrock upon which the last ten years of my life have been built; thank you and I love you.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

Modern military thinking divides war into strategic, operational, and tactical levels. This notion is attributed to Helmuth von Mulke, who led the Prussian, and eventually German, army through many successful wars in the 1860s and 1870s [1]. By this timeframe, warfare had changed significantly. The railway and improved road systems allowed larger armies to be mobilized and dispersed across a wider area; the telegraph allowed for centralized control of these dispersed forces practical; and improvements in agriculture and industry allowed the larger, more mobile forces to be sustained. Technological advances in weaponry such as artillery, machine guns, and aerial bombardment provided a necessity for troops to further disperse to minimize their lethality. Thus, by World War One, a single battle stretched across hundreds of miles of trenches to be fought over for four years; instead of a single field in which victory could be claimed within a single day [1].

As the twentieth century progressed, the mechanization of the infrantry and the advancement of aerial capability further stretched the area of military operations. Thus the edge of the battlefield was blurred. For the first time, firepower could be excercised by aircraft, artillery, or missiles outside the immediate area of conflict. Now, maneuver of troops and supply outside the area of conflict, have an immediate

impact on the area of conflict [1]. Neither the tactical nor the strategic levels encapsulated the requirements of this new theater commander; thus the operational level was added to bridge the gap between the two. These three levels of warfare allow causes and effects from every aspect of warfare to be better understood as a whole; despite its growing complexity. The Department of Defense [2] defines them as follows:

Strategic The level of war at which a nation determines national security objectives and guidance, then develops plans and policies for the use of national resources to achieve those objectives.

Operational The level of war at which campaigns and major operations are planned, conducted, and sustained to achieve strategic objectives within theaters or other operational areas.

Tactical The level of war at which battles and engagements are planned and executed to achieve military objectives assigned to tactical units or task forces.

The introduction of networked systems into warfare has only further blurred the edge of the battlefield. Cyberwarfare is occurring continuously across world-wide network connections, resulting in minor disruptions, website defacements, theft of national defense information, and intellectual property theft. This addition of cyber has introduced new challenges to traditional well-understood military concepts at the each level of warfare.

At the strategic level, cyber warfare is very much a contentious issue. To briefly illustrate this, in 1993, John Arquilla and David Ronfeldt wrote an article entitled

"Cyber War is Coming!" [3] whilst, more recently Thomas Rid wrote an article entitled "Cyber War Will Not Take Place" [4]. During this time frame, the Stuxnet attack managed to take control of physical infrastructure and disrupt Iranian nuclear ambitions [5] whilst the cyber domain also allowed the unstoppable publication of massive amounts of secret diplomatic cables and intelligence community secrets [6]. In chapter 2, I analyze the Department of Defense's inclusion of cyber as a separate and equal domain of warfare and discuss how the unique characteristics of cyberspace do not fit well into the molds of the traditional warfare domains of land, sea, air, and space. US policy makers can use the basis provided to create a cyber doctrine for the military that is also in line with the strategic objectives of the nation beyond military force, that embraces cyber's unique aspects while letting go of the ideational frameworks of the other domains.

At the operational level, the dependency of the United States, including the Department of Defense, on information systems and networking technologies places the nation at risk of a loss of confidentiality, integrity, and availability of its critical information resources, degrading its ability to complete the mission. To combat this threat, situational awareness in cyberspace is critical. Therefore, I analyzed each cyber situational awareness tool in the DoD's toolbox, discovering each systems strengths and weaknesses. While I cannot publish this analysis due to confidentiality concerns, in chapter 3 I introduce an operational framework for establishing situational awareness in cyberspace which was developed from the information discovered in the analysis.

The DoD has already begun integrating this framework into a new generation cyber situational awareness tool.

At the tactical level, the rise of social media has created a growing tension between military users personal needs and military operational security. Intelligence plays a vital role in the day-to-day tactical operations of our military. In chapter 4, I investigated the utility and feasibility of an adversary aggregating open source information to create possible intelligence targets. In doing so, I created an automated approach to data collection, analyzed the data through content analysis and applied machine learning techniques to learn as much as possible from the data collection. Military members found were then ranked based on vulnerability. In all, 1168 military members were deemed to be potential intelligence targets with 223 of those determined to be easily targeted by an adversary. These members are now open to exploitation in a variety of scenarios including being tagged while abroad (even in civilian attire) for surveillance or held for ransom, targeted for a cyber phishing attack, or killed by terrorists wishing to make a political statement here in the United States. Additional training, technical controls, military-civilian partnerships and legislation are discussed as possible remedies to ensure our military personnel are secure and remain vigilant in this time of war. Most importantly, I demonstrated that automated methods can be effective in discovering easily targetable personnel by an adversary.

Finally, Chapter 5 presents concluding thoughts and potential topics of research for the future of cyber warfare and how it fits into the current construct of military

thinking at the strategic, operational, and tactical levels.

CHAPTER 2

Unvalidated Input: Reconsidering Cyberspace as a Warfighting Domain

> Like everyone else who is or has been in a US military uniform, I think of cyber as a domain. It is now enshrined in the doctrine: land, sea, air, space, cyber...But the other domains are national, created by God, and this one is the creation of man. Man can actually change this geography, and anything that happens there actually creates a change in someone's physical space. Are these differences important enough to rethink our doctrine? [7]
>
> —General Michael V. Hayden, USAF, Retired
>
> Former Commander National Security Agency

2.1 Introduction

General Hayden, who served as director of both the National Security Agency and the Central Intelligence Agency through nearly all of the last decade, witnessed a tumultuous shift in US security priorities, process, and strategy prompted by the Al Qaeda attacks of 2001 and extensive military interventions that followed. In those positions, Hayden presided over a massive reorientation of military and intelligence

capabilities away from post-Cold War to counterinsurgency, counter-terrorism, and deterrence missions applied against ideological groups and a small set of rogue or rogue-ish states. But is he, along with everyone else who has served in uniform, correct that cyber is indeed a new domain, one distinct from the others, yet still holding sufficient analogs to the other domains to be considered such? This I consider in my pursuit of understanding what the domain of cyberspace is, how it differs from other domains, how conflict may be undertaken within it, and how militaries, including that of the United States, may seek to employ force in cyberspace. While Thomas Rid asserted, "Cyberwar will not happen," in his book's title, I make an ancillary argument that cyberspace may not be a distinct domain from all others, inasmuch as all conflict involves human interaction, but humanity is not a domain. Further, I offer that perhaps information and computing issues are better left embedded in the other domains as well as the set of interactions by which states communicate, cooperate, and disagree with one another below the threshold of conflict.

2.2 Cyber Does Not Equal Domain?

Any consideration of how the United States exerts military force in cyberspace requires at least brief consideration of what it is. The Defense Department defines cyberspace as, "A global domain within the information environment consisting of the interdependent networks of information technology infrastructures and resident data, including the Internet, telecommunications networks, computer systems, and

embedded processors and controllers"[8]. No doubt much staff work went into this description, however, this appears to define cyberspace as the sub-components of an information environment, the technical infrastructure. Definitions of cyberspace have been widely considered for the last several years, but the term has become more of a fixture because of the cyber security issue. 'Cyber' and cyber security, in many circles, are nearly synonymous. Entry of 'cyber' into the creative vocabulary can be attributed to Norbert Weiner through his thinking on cybernetics, a science of animal (including human) machine communication [9].

Cyberspace is a term first used to describe the interconnected digital venue in William Gibson's Neuromancer. "Cyberspace. A consensual hallucination experienced daily by billions of legitimate operators" [10] But the name of this consensual hallucination stuck and has come to describe what is likely the great marvel of our time, the computer and its connectedness to other computers on a global scale. To make clear my own perspectives, I generally agree that some sort of cyberspace may exist, just so long as I can firmly tie it back to its sci-fi roots and its birthplace at a connective point between information theory, zoology, and philosophy located by Wiener. What cyberspace has grown up to become is some summarizing term for the computers that have become internet-worked and largely ubiquitous; including the pprt[1] cognitive augmentation devices we each typically carry and call phones. This infrastructure has grown very important to military affairs, in the U.S. and many

[1]pprt is drawn from another work of fiction, with a setting in the not too distant future, and the term describes small, interconnected computers that seem to draw enormous attention from their owners [11].

other armed forces from those of Russia to the Islamic State movement.

But in the US defense establishment, the United States Air Force was first mover on cyber, by signaling its intent to invest heavily on operations in cyberspace in 2007, an act that would lead observers to question if it were to be the lead US service for cyberspace. Following it, the United States Department of Defense (DoD) has put considerable effort into developing forces designed to engage in operations within cyberspace [12]. This culminated with a declaration in 2011 of cyberspace's equal standing at the DoD as a domain of conflict, standing alongside land, sea, air and space as well as issuing policy for operations in cyberspace [13]. This is an incredible act, not to be brushed off as fad, as it declares the topography of computers and links between them on planet earth and above it as an area of emphasis and concern for US military activities. Each of the armed services holds significant experience in their respective domains, and from it doctrinal guidance regarding the use of force and capacity to exert influence by military forces has emerged.

Knowing the military organizes, trains, and equips within 'domains' of conflict, I offer this chapter as a critique on the declaration of cyberspace as a domain and doctrine of cyber warfare as it stands. I also wish to bridge thinking in computer and information science with that regarding international politics and strategy in an interdisciplinary stab at critiquing the viability of a national policy that does no less than declare the world's computers and the links between as topography for battle. No doubt, there is incredible contemporary interest in cyberspace (a phenomenon I

will describe more completely later) and its connection to politics. As the events of the Arab Awakening, the Stuxnet episode, and the still unfolding Snowden affair suggest, cyber issues are now a significant component of geopolitics, relevant to international public opinion, bilateral and multilateral diplomacy, and military operations. As a result, I argue that cyber matters, perhaps a great deal. Nonetheless, there are tough questions regarding how the United States has chosen to militarily define cyberspace and whether doctrine regarding it is tenable or even advantageous for U.S. standing.

Central in my thesis is whether cyberspace is not a domain as currently defined and whether the DoD should re-consider the model of cyber operations distinct from other domains. To me, this is a valid question, as the dividing line between digital technologies and military operations is blurry. This chapter considers what cyberspace is and how it may be construed as a distinct domain. It compares how cyber is different from other domains, and how operations within it compare to those with which are more familiar. Computational innovation was largely born of military need during the Second World War, but only with the rise of the Internet as a global communications and computing medium has it gained its distinct identity in military circles. To better understand this shift, this chapter begins with consideration of the concept of domain and military operations.

2.3 What is the Cyberspace Domain?

The DoD defines a domain as a "territory over which rule or control is exercised" [14]. A domain is an operational environment representing a physical manifestation where military operations may be conducted. Land, sea, air, and space, the traditional warfare domains, are areas in which services cooperate in joint operations, but with the exception of space, each clearly aligned to the core missions of one service [2]. Now, cyberspace has been accepted as unique from the others; defined as, "A global domain within the information environment consisting of the interdependent networks of information technology infrastructures and resident data, including the Internet, telecommunications networks, computer systems, and embedded processors and controllers" [2].

While treating cyberspace as a domain might establish a foundation from which to understand and define its place in military operations, how control over it may be asserted and to what degree such control is asserted over the function of computers or extends to the ideas and images they may be used to convey. Although implementation of policy in land, sea, and air domains has an information dimension applied through capacity for communications, intelligence, disinformation, and persuasion. For militaries, information has two primary applications: command and control, and intelligence. While I understand cyberspace to be some collection of computers and information spread across the planet, "rarely has something been so important and so talked about with less clarity and less understanding than this phenomenon [7]. Sim-

plistically, I assert that it is man-made; easily and constantly replicated; composed of physical, syntactic, and semantic layers; and the costs of entry to it are relatively low [15].

Cyberspace is the contemporary medium for military communications. Before the advent of computing and networking technologies, militaries used carrier pigeons, dispatch riders, and signal flags to communicate and employed agents or observers to gather intelligence for specific missions [16]. Telegraphy, radio, and other analog communications technologies were gradually adopted for the communications mission, along with ancillary intelligence function aimed at the analog electronic communications of adversaries and enemies. Pre-digital assets served communications functions, but were generally organized into functional signals branches within the services, something akin to engineer or quartermaster corps. The revolution in information and computing technologies (ICTs), however, contributed in the late 1990s to a belief that a digitally driven revolution in military affairs (RMA) was underway[2].

Accepting the confluence of ICT and RMA, DoD's 2001 doctrinal document, "Joint Operations" identified five war-fighting domains, adding an information domain [17]. This inclusion sparked debate among military thinkers about how information fit into the current operational environment. Previous clarity on commonly accepted operational roles and functions became blurred, including which mission

[2]For information on how the cyber use in military operations extends beyond command, control and communications, check out the unclassified summary of Air and Sea Battle Concept from the Office of the Secretary of Defense at http://www.defense.gov/pubs/ASB-ConceptImplementation-Summary-May-2013.pdf

sets would fall under this new information domain. Arguments were made for missions ranging from electronic warfare, leaflet dropping missions, the nations satellite constellations, and even the airborne laser project to be included in the new domain [18, 12]. Unable to reach doctrinal consensus, information in the 2006 version of "Joint Operations" was re-characterized from a war-fighting domain to an 'environment' [19].

However, this change did not resolve fundamental disagreements. In 2011, the United States Department of Defense Strategy for Operating in Cyberspace repackaged the information concept and formally declared cyberspace as an "operational domain to organize, train, and equip so DoD can take full advantage of cyberspace's potential in its military, intelligence, and business operations" [20]. What remained unclear is what war in cyberspace looks like. Libicki saw two forms: cyber warfare and cyber war. "Cyber warfare is about the conduct of war, carried out inevitably to further the performance of combat in the physical domain. Cyber war is undertaken to affect the will of the adversary directly" [21]. The cyber attack purported by Clarke and Knake by the Israelis against Syria's integrated air defense system, possibly undertaken via an advanced electronically scanned array (AESA) airborne radar, falls nicely into Libicki's category of cyber warfare as an enabler for an air strike against a strategic target, in this case a Syrian nuclear facility [22]. Stuxnet would stand as an example of cyber war, as might the denial-of-service attacks launched by the Syrian Electronic Army against the NY Times and other news outlets when the US

leadership was weighing the options on punitive strikes against the Assad regime.

While the security of information and computing systems will likely be part of any future war between sophisticated opponents, I consider Bruce Schneier's advice that cyber operations may not be divorced from broader conflict of a kinetic nature [23]. When influencing the will of an adversary in contemporary conflict there are a variety of meanings. To thinkers on air power, discussion regarding the will of the adversary begets a conversation on the much-debated value of strategic bombing during the Second World War. Hitler's expectations of British capitulation at the hands of the Luftwaffe in 1940 were informed by the dramatic effect of bombing Guernica and Rotterdam, but when Bomber Command burned Hamburg and Cologne, Hitler was no more eager to sue for peace than Churchill had been in 1940. This casts doubt that a cyber campaign would drive an enemy state to capitulation.

But there is a second facet of cyber war that I must consider before moving onto discussion of cyberspace itself. Today's DoD has been shaped by more than a decade of US counter-insurgency operations in Iraq and Afghanistan, coupled with counter-terror operations from the Maghreb to the Philippines. In addition, it is forced to cope with adversary states in the global system, such as Russia and China. Cyber tools may be employed to counter Jihadist terror recruitment or block the regional ambitions of states, often with nuclear arsenals or the aspiration to build them, but these are just tools, not necessarily a new form of distinct, major military power[3].

[3]Additionally, in differentiating cyber war and warfare, it is necessary to note foreign interpretations of information security versus cyber security. In diplomatic discussions with Russia and

2.4 How is Cyberspace Different?

As General Hayden reminds us, cyberspace is a man-made environment. Another man-made creation is the city and contemporary militaries often operate in densely populated urban areas that are a terrain as foreboding to military operations as any desert or jungle. Stalingrad, Hue and Fallujah stand as exemplars. However, there is something mutable about cyberspace. It is dynamic and complex, and at times its structures seem fragile, as when a Pakistani edict against the Google YouTube video service became a major outage as routing data for the Boundary Gateway Protocol (BGP) meant to block access in Pakistan soon reached across the globe. It is an example for Nye's point, that, "Mountains and oceans are hard to move, but portions of cyberspace can be turned on and off by throwing a switch" [24]. Furthermore, the architectural components of cyberspace may be fairly easy to map, navigate, open or close, but when I begin to consider it as an ideational area, concepts of control-moving from a syntactical to semantic understanding of cyberspacemuddy the waters in a manner analogous to the difference between traditional military operations and unconventional ones such as covert action or propaganda campaigns. Therefore, it is difficult to apply the tenets of warfare that exist in the other physical realms, especially those geared toward traditional conflict. Stytz and Banks assert, "Achieving

---

China through the UN GGE and other meetings, the Russian and Chinese interpretation of cyber security generally regards the protection of information systems from disruption, essentially breakdowns in confidentiality, availability, integrity, non-repudiation. Information security, on the other hand, represents state control of information, across the press, social media, and other vehicles of dissemination, all in digital form.

global cyber superiority or global cyber control by any organization is no longer possible " [25]. But with a broad view of what composes cyberspace, as both technologies and information, I visit the hard problem of how to draw a clean dividing line between targets in the syntactic layer from those within the semantic one.

Nonetheless, the computational underpinnings of cyberspace continue to evolve. Moore's Law, even if it again is in danger of breaking, has brought with it immense computational power at prices affordable to all corners of the globe. Cyberspace is the computing hardware, personal computers, tablets, smart phones, data centers, switches, routers, and other gear for the processing, storing and transmitting data, so much of it ephemera. Remarkable for the feat of global buy-in, the ubiquitous Transmission Control Protocol-Internet Protocol (TCP/IP) standards interconnect billions of devices up to major communications backbones and back down to recipients. This is the architecture of computing that composes cyberspace. I compare the most recent of the classical domains (as war has not yet visited space with the significance of air power) with how cyberspace appears (See Table 2.1).

One of the definitional challenges for understanding cyber conflict is the label 'cyber attack,' a term which is generally accepted in computer security to be an incident in which a malicious actor in some way compromises or damages a system or its contents. Although Hathaway et. al. argue for a more narrow national security-oriented definition [26], almost any malicious act involving computers is often labeled as an attack. The term has stuck. When an attack is noticed, network administrators,

Table 2.1 : Comparison of Air vs. Cyber Domains

	Air	*Cyber*
Environment	Natural, constant	Rapidly changing, man-made
Primary Roles	Aerial reconnaissance, bombardment and close air support	Intelligence gathering and disruption of enemy decision making process
Technology	Expensive to enter, years to traverse acquisition cycle	Cheap to enter, constant technological advancement
Recon	Fly into enemy territory, locate assets and report	Ascertain input from any of the 8.7 billion devices connected to Internet
Attack	Air to air combat or bombing enemy target; lives at risk; aircraft and crew hard to replace; superiority possible	Opposing cyber forces do not meet head to head; great uncertainty in outcomes and even mapping of resources to results; gains may be informational or rhetorical as much as physical
Defense	National borders demark boundaries; stop all attacks on U.S. assets	DoD firewall demarks boundaries, do not protect civilian assets; strategy to ensure mission success not to stop every attack
Battle Damage Assessment	Overhead imagery, sensors, and other intelligence sources employed for assessment	Compromise of hosts or systems potentially easy, but impact of attacks difficult to ascertain
Attribution	Mission recording and weapons dispensed provide details	Forensics informed by motive and perceived adversary and state capacity
Proportionate Response	UN Charter, Hague and Geneva Conventions, and related treaties; standard ROEs state if fired upon, can fire back in kind	International norms in development, but no clear definition or national threshold determined on what constitutes an act of aggression
Deterrence	Fear of retaliation based on previous show of force	Deterrence by denial through hardening of defenses

in seconds, can change the logical landscape of the network as seen by an attacker. Operating systems and applications can be virtualized and reset to previously known safe points. Information may be moved to safe locales for preservation and eventual recovery.

An example of this resilient recovery mechanism emerged from the 2008 Russia-Georgia conflict. Three days before the commencement of kinetic hostilities, Russian hackers[4] launched distributed denial of service and defacement attacks on Georgian websites. Georgian officials responded by hosting the blocked content on the infrastructure of US corporations, such as Google. The attacks continued but efficacy fell off [27]. This lends evidence that the possibility of collapse of cyberspace is grossly overstated. While overwhelming large systems is still possible, configuration modifications and backup capacity assures that these systems stand a solid chance of recovery from failure.

While thoroughly crippling systems may be hard, participation in cyber conflict requires me to consider the cost of bringing cyber weapons into conflicts [28]. While becoming a significant actor in the land domain usually carries a hefty tab, air and sea power is even pricier today. Building aircraft carriers, nuclear submarines, stealth fighter aircraft and main battle tanks is an activity for nation states. These platforms are the pillars of a form of conflict, international warfare that was rendered largely

[4]In security lexicon, a hacker is one who uses programming skills to gain illegal access to a computer network or file (American Heritage Dictionary). The concept of hacking, however, has a broader meaning to computer culture, which may characterize the term as "an appropriate application of ingenuity." See: The Jargon File. http://www.catb.org/jargon/html/index.html.

unthinkable because of nuclear deterrence. Those countries with the means to create significant armies, navies or air forces have strived to develop nuclear arsenals or forge security agreements with those powers willing to stretch their nuclear umbrella. This has not spelled the end of conflict, as the actions of Al Qaeda, Hezbollah, the FARC, the Lord's Resistance Army, and any number of active insurgent groups continue around the globe. Symmetric battles such as Austerlitz, Trafalgar, Gettysburg, Jutland, Midway, and Kursk have largely given way to asymmetric fights falling under the labels of terrorism and insurgency.

During the 20th Century, the United States largely achieved dominance in the land, sea and air domains by spending heavily on research, development, and acquisition of the most technologically advanced weapons platforms. With the end of the Soviet Union, attempts to outpace the US and its Western allies with more tanks, warships and aircraft ended as well [12, 24]l. But in cyberspace, participating in malicious activity is cheap. For $1,000, a botnet of 10,000 infected computers all located within the United States can be bought [29]. Too expensive? Freely available tools, such as metasploit, can be used to detect computers ripe for compromise. Zero-day exploits can be purchased in the malware marketplace for as little as $5,000 for Adobe Acrobat Reader or $60,000 for Windows, Firefox, or Microsoft Word [30, 31]. Compromise of systems may be accomplished by delivery of a well-crafted targeted e-mail (spear phishing) that once opened, leads a user to download malware directly to their machines, bypassing any and all security mechanisms in the process. Such

easy access to infected machines and a large body of ready to use system exploits may make it difficult for any nation to exert dominance in cyberspace.

In addition, the Internet was designed holding paramount the concepts of reliability and free flow of data, through common protocols. However, in wiring together the globe, a network built to support the activities of highly trusting actors, principally academics, has been delivered to a far broader and less trustworthy crowd. Every fraudulent email and online scam stands as testimony to support this argument. Although law enforcement and legal systems are constrained by physical borders, adversaries are not. Distance does not matter; the enemy is located just on the other side of a wire. Friend and foe alike are being connected into the same network. Even weakly governed developing countries are being connected using fiber-optic cables, opening new safe havens for cyber adversaries as any coffee shop with a WiFi signal can be used to launch an attack [32]. There is no dominating an enemy that can use a few hundred dollars worth of equipment and from anywhere affect the process of a commercial or government target. The means of cyber conflict are cheap equipment and knowledge. With this the tableau of potential conflict, I must consider what doctrinal thinking holds relevance.

2.5 Do Traditional War Fighting Concepts Work?

To claim cyberspace as a domain promotes the use of war fighting concepts from the earlier domains of land, sea, and air. From these domains, military leadership draws

upon 'cybered' interpretations [] of activities including intelligence gathering, attack, defense, attribution, proportionate response and deterrence. Although classical military theory remains relevant in contemporary military doctrine, land, sea, and air have also seen significant more recent domain-specific theoretical contributions as well as those combining multiple domains. Liddell-Hart, Mahan, and Douhet each shaped how military power would be exercised on land, at sea, and in the air, respectively during the last century, however, Libicki reminds thinkers on cyber conflict, that, "If the Owl of Minerva flies at dusk, in cyberspace the sun is just above the yardarm; the information revolution is hardly a done deal" [21]. With this the case, I ask if US policy to build a distinct Cyber Command and etch out the doctrinal space for cyber alongside the three traditional domains and the one in which navigation, communications, and overhead surveillance satellites toil for a growing club of nations. I begin with the precursor activity to military action: intelligence.

2.5.1 Intelligence Gathering

Intelligence[5] is "the process by which specific types of information important to national security are requested, collected, analyzed, and provided to policymakers" [33]. Open intelligence gathering in cyberspace uses sources such as public websites, chat rooms, bulletin boards, blogs, forums, and discussion groups to obtain very tacti-

[5]Intelligence is the product resulting from the collection, processing, integration, evaluation, analysis, and interpretation of available information concerning foreign nations, hostile or potentially hostile forces or elements, or areas of actual or potential operations. The term is also applied to the activity which results in the product and to the organizations engaged in such activity (JP 2-0).

cal information on potential enemies. Due to the use of such sources by extremists for communications, recruitment, fundraising, and spreading propaganda, the United States has made extensive use of open source and cyber intelligence gathering techniques in its ongoing counter-terrorism and counterinsurgency operations [34].

Espionage[6] involves clandestinely obtaining information considered confidential without permission of the rightful owner. Clandestine intelligence operations have been a part of statecraft since Rahab's work as an agent of the Hebrews at Jericho. Today, cyber means permit massive wholesale espionage, extending even to design plans for newest US fighter [35]. Cyber operations have according to former US defense official William Lynn, permitted the theft, "Each year, an amount of intellectual property many times larger than all of the intellectual property contained in the Library of Congress is stolen from networks maintained by US businesses, universities, and government agencies " [36].

Reconnaissance[7] is an exploratory survey to obtain military information, particularly of the makeup of an enemy force. In the traditional domains, this meant sending scouts, usually into enemy territory to observe and report on enemy dispositions and resources. Reconnaissance is also integral to offensive cyberspace operations, as it is the logical mapping process that prepares an attacker to successfully exploit a

[6]Espionage is the act of obtaining, delivering, transmitting, communication, or receiving information about the national defense with an intent, or reason to believe, that the information may be used in the injury of the United States or to the advantage of any foreign nation (JP 2-01.2).

[7]Reconnaissance is defined as a mission undertaken to obtain, by visual observation or other detection methods, information about the activities and resources of an enemy or adversary, or to secure data concerning the meteorological, hydrographic, or geographical characteristics of a particular area (JP 2-0).

targeted computing device or network.

In US doctrine, these intelligence operations in cyberspace are often described under the term Computer Network Exploitation (CNE), which is defined as the "enabling operations and intelligence collection capabilities conducted through the use of computer networks to gather data about target or adversary automated information systems or network" [37]. Of intelligence in general, Keegan argues that intelligence is but a small component in the overall calculus of conflict. "Knowledge, the conventional wisdom has it, is power. Foreknowledge is no protection against disaster. Even real-time intelligence is never real enough. Only force finally counts" [38]. Kahn also views intelligence as constrained to optimization resources as an auxiliary function in warfare [39]. The facts that computers now store critical information and that the Internet is an important vehicle for collection do not fundamentally re-make the role of espionage for military services. That cyber intelligence activities may employ some subset of the more than 8 billion Internet connected devices on the planet indicates a revolutionary growth in sensing and collection capabilities [40], but does not provide adequate evidence that a cyber domain need exist today. Computing was prompted ahead to serve the signals intelligence and cryptanalysis capabilities of that served the Allies in the dark days before Alamein and Midway.

2.5.2 Attack

A doctrinal fixture of the DoD, Computer Network Attack (CNA) is the "use of computer networks to disrupt, deny, degrade, manipulate or destroy information resident in the target information system or computer networks, or the systems/networks themselves" [37]. Planning and preparing for an attack may take weeks or months to examine the logical structure of the target and determine its vulnerabilities. Once launched, the strike may be over in matter of seconds. Therefore, in many cases, it is not realistic to assume a defender can recognize the attack and react while the attack is in progress. This is why defense-in-depth is still considered the most prominent philosophy in repelling an attack; harden the most damaging avenues for attack in advance, search for anomalous behavior, then react as quickly as possible to mitigate and remediate the effects of an attack once it occurs. Of this process, Libicki observed, "defending the network is not so much to maneuver better or apply more firepower in cyberspace but to change the particular features of one's own portion of cyberspace itself so that it is less tolerant of attack" [41].

Aside from being an avenue for collecting intelligence as discussed above, offensive cyber operations serve two purposes: deception and disruption. The goal of deception is to "mislead an enemy by manipulation, distortion, or falsification of evidence to induce the enemy to react in a manner prejudicial to the enemy's interests" [2]. This is done in many ways. Attacks are easily concealed by routing through the TOR anonymous network [42, 43]. Trojan horses camouflage themselves as benign pro-

grams. Phishing e-mails and websites masquerade as legitimate, delivering malicious payloads to computer systems.

Hollywood depicts cyber war with hackers breaking into computer systems controlling and disrupting phenomena in the physical world [24]. Leon Panetta delivered a stark warning in 2012. "We could face a cyber attack that could be the equivalent of Pearl Harbor" [44]. To date the majority of disruptive cyber attacks benefitted one of the other domains: changing the red dots to blue dots on an enemy's command and control display or making friendly aircraft disappear on an air defense radar (i.e. Israel's 2007 airstrike on Syria's nuclear reactor) [45]. Employing CNA against another cyber force does not appear to be a quick avenue to eroding capability as an adversary's cyber capabilities are derived from its hackers, their cyber tools, and the intelligence they collect none of which can be destroyed entirely through cyber means. "Since hackers need only an arbitrary computer and one network connection, it is not clear that even a physical attack could destroy a state's cyber attack capabilities" [18]. An opponent's computer system can be replaced for a few hundred dollars if corrupted or destroyed via cyber means.

The United States cannot expect to substantially diminish an adversary's offensive cyber capability through offensive cyber means (or necessarily even by kinetic means) [46, 47, 48]. One USAF officer argued, "if we engage in a cyber war with inferior forces, we cannot depend on superior tactics to outmaneuver an opponent, inflict greater losses, and turn the tide" [12]. But could the same not be said of counterinsurgency

operations? For such campaigns of this sort a well-educated, trained, and capable force must be able to outlast the adversary [49, 50].

As with early air doctrine, cyber operations are currently viewed as a means to soften the enemy before launching kinetic operations [51]. However, contrary to the examples of Billy Mitchell and Hap Arnold of the early days of airpower, the idea that cyber alone can win future wars, has not gained credibility. The United States can achieve advantages through cyber attacks, such as creating a window of opportunity to steal data or launch a kinetic attack, but without the ability to destroy an enemy's cyber resources, long-term control or superiority cannot be achieved. Thus, offensive cyber operations today lack the decisive characteristics of land, sea, and air in defeating enemy forces or deterring aggression. Cyber attack largely appears to be a sort of psychological operations tool or a piece of command and control warfare than that of a new cyber war-fighting domain.

2.5.3 Defense

When the need to defend computer systems arose, the activity was labeled information assurance. Information Assurance (IA) is a set of "measures that protect and defend information and information systems by ensuring their availability, integrity, authentication, confidentiality, and non-repudiation" [2]. William Lynn described defensive cyber strategy as having five main elements: "Develop an organization construct for training, equipping, and commanding cyber defense forces; employ layered

protections with a strong core of active defenses; use military capabilities to support other departments efforts to secure the networks that run the United States critical infrastructures; build collective defenses with US allies; and invest in the rapid development of additional cyber defense capabilities" [36]. Collectively, these elements point to the goal of being able to complete a wide range of missions assuming network degradation.

Although cyber attacks may develop over time to look more like Stuxnet, but most are acts of espionage or theft [52]. Military networks are defended to ensure the availability, integrity, and confidentiality of the data resident on and passing through them. Although the DoD is doing everything it can to properly defend military networks, the extensive record of successful attacks against DoD information systems and those upon which it relies demonstrate cyberspace superiority remains out of reach. The DoD no longer attempts to stop all attacks by securing every piece of the network but now assumes an attacker is capable of getting into the network (or is already there) and that "operating with a presumption of breach will require DoD to be agile and resilient, focusing its efforts on mission assurance and the preservation of critical operating capability" [20]. This presumption of breach runs at odds with any control or dominance strategy as it implies that an adversary may be able to routinely manipulate or outpace the OODA loop of US or allied forces. Defending military computing networks becomes much more about risk management than warfare.

Another aspect of the traditional warfare domains that does not correspond in

cyberspace is in assignment of defensive roles and responsibilities. US policy has defined some roles for military forces in cyber defense, but unlike land, sea, and air, if US civilian assets are attacked via cyber means, the DoD does not hold the primary defense role unless requested by the Secretary of Homeland Security [13, 53]. Lin posits two legal responses to a cyber attack on a non-government entity: "First, a private actor can take measures within its organizational boundaries to strengthen its defensive posture; second, it can seek the assistance of law enforcement to investigate and take action against the threat" [54]. Neither of which may immediately stop the pain of a cyber attack. Reducing the military's role in cyberspace to chasing illegal bits only when they cross a military firewall severely limits the possibility of defending, much less dominating, cyberspace. With the private and public sectors together now forming the front line of any twenty-first century war, how can a civilian-owned entity be a war-fighting domain?

A more applicable model for cyber security may be that of civil defense in policy on defending information resources, but there is considerable reluctance for corporate actors to throw open their networks to monitoring by government entities [55]. A difference between military defensive roles and other security missions exists [56]. There are important questions about the balance of an offensive edge with defensive capabilities, including the role of cyber security firms with attack back capabilities and the market in publicly unknown zero-day cyber vulnerabilities. Such issues make cyberspace look less like a domain for the DoD to dominate in time of war than an

ecosystem or marketplace in which competition may eclipse actual conflict.

2.5.4 Battle Damage Assessment (BDA) and Attribution

Following an attack, both the aggressor and defender must to determine the extent of the damage that occurred and the attacked nation must determine who is responsible[8]. Consider the following scenario:

A nation wants to delay or disrupt the process of enriching Uranium by an enemy state. The nuclear enrichment site is well hidden and protected by being buried deep into a mountainside. Assume there are two options: either drop bombs on the entrances of the enrichment facility or use a cyber weapon such as Stuxnet to cause the centrifuges to over speed.

In the airstrike, a GPS-guided weapon is placed directly on target at the specified time. Overhead imagery can see the explosion occur and continue to monitor enemy cleanup efforts. The facility itself may not have been destroyed, but confirmation can be provided that the target has been hit and an approximation can be given for recovery. From the enemy's perspective, it is fairly easy to determine the extent of the damage by running diagnostics on the machinery and previous knowledge of construction. It is also possible to determine who launched the attack based on description of the aircraft and the technological puzzle recovered from the bomb itself.

If the attack were carried out in cyberspace, there are no visual cues as in the

[8]Battle Damage Assessment if the estimate of damage composed of physical and functional damage assessment, as well as target system assessment, resulting from the application of lethal or nonlethal military force (JP 3-0).

air strike, making it hard to determine if the cyber target was indeed the one or if the cyber weapon actually disabled the enemy's enrichment facility. The common techniques to overcome this uncertainty is to listen for inactivity from the target or to capture communications sent through other means confirming the intended disruption had occurred [12]. Yet, if the target is not the intended one, these forms of BDA have little meaning.

If correctly targeted, the enemy can easily mask damage or generate false effects such as appearing to slow down or disappear from the Internet. If the target is air gapped from the Internet, as in the Iran nuclear facility, success may depend on an enemy error, as discussed above. If an operational commander cannot trust cyber battle damage assessment, they would be much less likely to use such weapons in a primary role in the conflict. Cyber would thus be relegated to a secondary role that, when used, its effects must be double checked through one of the four traditional domains.

Assume the attack was successful and the enrichment facility's centrifuge malfunctioned and broke down. From the enemy's perspective, they would run the same diagnostics to determine what machinery is broken and needs to be fixed. However, since the cyber attack falsified safety and security readings on the infected machines, determining the root cause would be difficult. The deception causes the enemy to live under a cloud of fear and suspicion, never knowing if all malware is out of the system. Without replacing all software and hardware in the process, they may not be

able to trust their system again. In addition, unless someone claims responsibility for the attack or leaves signatures in the discovered code, attribution becomes a guessing game of perceived intent and expected adversary capabilities.

2.5.5 Proportionate Response

Considerable effort has been expended in understanding the concept of jus ad bellum, the right to war, in cyberspace [57]. Legal and ethical issues arise if someone fires a cyber weapon at US forces or interests, but when and how can the United States fire back? In traditional warfare, this is the standard operating procedure, however not the case in cyberspace. When a kinetic attack takes place on friendly forces, all avenues of response are available to ensure a proportionate response. The President asserted that, "when warranted, the United States will respond to hostile acts in cyberspace as we would to any other threat to our country," including, "the right to use all necessary means" [58]. However, cyber retaliation becomes problematic when considering the law of armed conflict's (LOAC) tenet of proportionality. LOAC was developed to cope with traditional kinetic warfare. How it applies to cyber conflict in a specific instance is uncertain. Small attacks of benign intent can have immense consequences while massive attacks of aggressive intent can have little to no consequences if poorly executed or properly defended. Without attribution and determination of motive, it is hard to determine the best course of action, especially responding to bits with bullets.

While norms in cyberspace (see discussion below) are currently being considered, there have been no clear guidelines set forth describing what an act of war, the proverbial 'cyber red line' would be. Current literature has spread characterization of cyber war (and warfare) from something that will never happen all the way to the sky is already falling [59, 3, 27, 4]. The reality is likely somewhere in the middle, and as stated before, most attacks are minor disruptions or stealthy malware designed for espionage.

Drafters of international law did not envision cyber capabilities and current law reflects this shortcoming, however, the United Nations Charter, Hague and Geneva Conventions, and related treaties are the primary instruments with which to assess acts of war. The UN defines aggression in Article 1 of the UN General Assembly Resolution 3314 as "the use of armed force by a state against the sovereignty, territorial integrity, or political independence of another state" [60]. Discussions convened by the UN in the Groups of Governmental Experts continue to stall in getting past differences on acts of aggression regarding information content or information infrastructure [61].

Where international accord does exist is in cybercrime. The Budapest Convention on Cybercrime was drafted to provide a common criminal policy aimed at protection against cybercrime. Forty nations have ratified it, agreeing to outlaw infringements of copyright, computer-related fraud, child pornography and violations of network security, and also agreed on electronic evidence sharing [62]. With only forty states

agreeing to the treaty (non-signatories include China, Russia, India, and Brazil) and with only basic definitions of what constitutes a cyber crime much less an act of cyber war, the Budapest Convention remains a commitment by a small club of countries on something outside the military sphere [63, 64].

A further step in international code on cyber conflict was NATO's Tallinn Manual, a non-binding document discussing the applicability of international humanitarian law and the concepts of a just war to cyber conflicts. It defined cyber attack as, "a cyber operation, whether offensive or defensive, that is reasonably expected to cause injury or death to persons or damage or destruction to objects" [65]. The Tallinn Manual's threshold is set at an operations kinetic consequences. By this definition, the denial of service and defacement attacks in Estonia and Georgia would not be considered cyber attacks, as physical harm to persons or objects did not occur. Interestingly, the authors of the manual could not agree on whether Stuxnet constituted an act of war. In discussing this point, Rid stated "if they cannot agree on whether the most sophisticated attack that ever happened falls within the realm of their own document, then they're basically talking about a class of events above Stuxnet, and that class is empty" [4].

2.5.6 Deterrence

In international relations, deterrence has traditionally been achieved through force or the threat of force. Since the Second World War, nations have developed nuclear

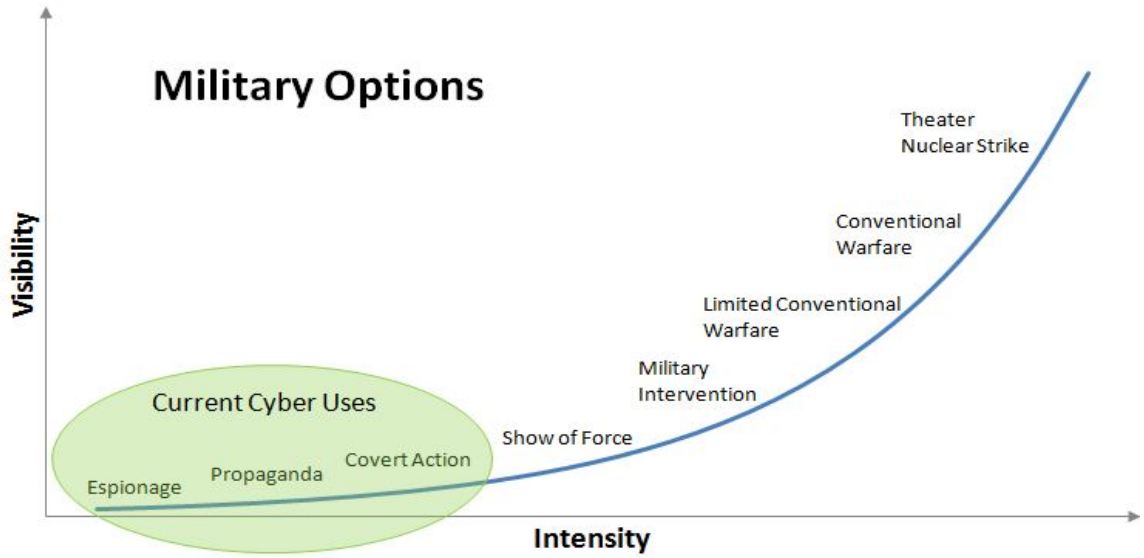


Figure 2.1 : Military Options: Intensity vs Visibility

arsenals to deter another full-scale global war for fear of nuclear annihilation. US conventional operations for the last two decades demonstrate a pronounced over-match capability in once contested areas including air combat and naval warfare. When the US-led coalition invaded Iraq in 2003, the Iraqi Air Force buried their aircraft rather than attempt to contest air supremacy [66]. But does this over-match extend to cyber operations? In 2007, Estonia found itself at the receiving end of a distributed denial of service (DDoS) attack for its decision to relocate a Soviet war memorial. The DDoS attack emanated from roughly 85,000 hijacked computers located in various countries including the United States and Canada. Analysts found postings online indicating Russian hackers were involved, but since the attack, no state has taken credit and no response has been issued [67, 48]. The political events surrounding the attacks and their sophistication provide evidence that the Russian

government must have been involved, something Moscow denies [67]. Despite all manner of scenarios for debilitating cyber attacks show of force in cyberspace has been very limited. Stuxnet, Shamoon (the cyber attack against Saudi Aramco) and attacks aimed at financial and media targets ostensibly launched by Iran or Syria represent the major political incidents.

Another matter confounding the construction of deterrence is that direct attribution is often hard, but not always impossible. With little to no fear of attribution, the only real deterrence for a nation is self-deterrence. Science journalist John Horgan offers a powerful reminder of how political leaders in the United States should consider their country's strength in this area with a revisit from one of the world's older geopolitical primers, Thucydides' *History of the Peloponnesian War*. Although the United States has employed cyber tools, often veiled under the protection of covert action, it cannot expect them to remain benign to its interests. The United States can expect other powers will employ a cyber weapon against it, its allies, or targets of strategic interest. Horgan suggests,

> We should consider the fate of Athens, which at the beginning of the Peloponnesian War was Greece's major power. Athenian soldiers eventually overran Melos, killed all the men and enslaved the women and children. But just as the Melians had predicted, the cruelty and arrogance of Athens aroused opposition against it. Sparta and its allies eventually crushed Athens, which never regained its former glory [68].

Demonstrating a capability in cyberspace may force an escalation in cyber capabilities. Many consider Stuxnet to be the starting pistol for an unregulated arms race in cyberspace [5]. On the other hand, demonstrating a cyber weapon may mean revealing a capability, which allows enemy states to develop countermeasures. Often, new variants of malicious code originate from another virus which has been modified. The biggest families like 'Zeus' have as many as 70,000 variants [69]. Even Stuxnet has siblings, Duqu and Flame (which, unlike Stuxnet, were not designed to disrupt facilities, but were conventional espionage tools) [5].

In the traditional domains, capabilities can be revealed, potentially forcing an arms race, but without revealing the technology or security vulnerability that has been exploited. And in invoking the topic of deterrence, I make a final argument on past models for deterrence by refuting the corollary from nuclear deterrence to some form of cyber deterrence. Clarke and Andreasen make a valid point, that, "Excessive rhetoric on the threat of cyberwar from the United States and blurring the distinction between cyber and nuclear attacks just makes progress toward cyberpeace more difficult" [70]. With such difficulty in asserting the capability to deter by cyber means or in cyberspace, I must ask again if it really can stand as a distinct warfighting domain.

2.6 Discussion

The United States relies on networked computing for all manner of economic, social, and civic activity. However, cyberspace also presents potential adversaries with an avenue to overcome the overwhelming advantage currently enjoyed by the United States in conventional military power. Cyber attacks can have instantaneous impact, are difficult to trace, and have the potential to produce significant harm to the United States and its interests. At least in part, this new set of risks has driven the Pentagon to consider cyberspace as a war-fighting domain.

In determining the correct doctrinal model for cyberspace, the ideological legacies of the traditional domains influence strategic thinking regarding problems, innovations, and strategies. Leaders tend to dismiss information that does not fit well with their preconceived notions and give excessive weight to information that is consistent [71]. With no examples of cyber war from which lessons-learned can be drawn, policy makers cling to notions of the other domains, attempting to force them onto this new arena. Returning to Hayden, I agree that, "casually applying well-known concepts from physical space like deterrence, where attribution is assumed, to cyberspace where attribution is frequently the problem, is a recipe for failure" [7]. As a war-fighting environment, the ultra high-speed, fluid, and omnipresent nature of cyberspace makes it fundamentally different from the traditional physical domains. So before imposing past tenets and terminology onto this new field, the question should be asked whether they are pertinent and to what degree.

Despite the trouble with antecedents and useful metaphors, the events surrounding Estonia, Georgia, and the Stuxnet worm have brought the question of cyber warfare to the forefront of world politics. Unfortunately, without agreed upon norms of acceptable behavior in cyberspace, nations for now are on their own in dealing with cyber attacks. US policy makers must create a cyber doctrine for the military that is also in line with the strategic objectives of the nation beyond its capacity for use of military force, hopefully embracing its unique aspects while letting go of the ideational frameworks of the other domains.

With approximately six drone strikes per week (or 300 per year) across Iraq, Yemen, Pakistan, and Afghanistan in recent years, I am able to convincingly argue that the United States is in a state of war in the air domain [72]. If cyberspace is a war-fighting domain, with over 46,000 cyber incidents reported to US-CERT in 2012 (22,145 by federal agencies), is the United States currently at war in cyberspace and will the United States be perpetually at war there [73]? If so, the urgent need for meaningful, domain-specific doctrine has never been greater. But perhaps we should not seek to so quickly characterize these incidents, many the product of poorly implemented software code, in a framework of war and peace. Most of the activity labeled as cyber security falls somewhere between these two poles, and can be undertaken for motives from criminal greed to issue-specific idealism.

There is also the matter that cyber is not tied, as is the case in land, sea, and air, to a full military service. The US Army Signal Corps issued its first solicitation for

construction of a flying machine in 1907, but an independent air force would not be stood for another 40 years, built upon an army air force composed of 2.4 million men and 80,000 aircraft at its peak [74]. The early adopter of the independent air arm concept, Britain, formed the Royal Air Force in 1918 [74]. There are plenty of reasons for not building a cyber service, with the U.S.'s successful shift to joint operations chief among them. Furthermore, the co-location of Cyber Command with the headquarters of the National Security Agency begs the question of how the information environment in which the cyber domain resides isn't just an intelligence function.

Finally, I must continue to ask, "How real is cyberspace?" On this question rests the validity of concepts regarding its mastery, control, and domination as well as its governance. Returning to the science fiction from which it is rhetorically connected, there is something real about it, but also realize that limiting cyberspace to the sum total of its computational parts rather than something that includes the semantic components of the 'information environment' referenced in the Defense Department definition isn't adequate. Instead, I like to see cyberspace as a system, in which all manner of interactions can take place, rather than a physical backdrop upon which events occur. While I remain unconvinced of the need for a distinct cyber domain for military operations, I accept that it may be possible. Ray Bradbury is associated with the quote, "Anything you dream is fiction, and anything you accomplish is science, the whole history of mankind is nothing but science fiction" [75]. If there is an irrefutable proof for the existence of cyberspace, then I will withdraw our arguments regarding

its validity as a distinct military domain.

CHAPTER 3

Operational Data Classes for Establishing Situational Awareness in Cyberspace

3.1 Introduction

The critical computer networks of the United States play a key role in our everyday lives, controlling the nation's energy, transportation, and financial systems. As such, the Department of Defense (DoD) has built operational dependency on its information systems and their associated networks. Disruption of these networks would have significantly damaging effects on the United States' ability to operate and defend itself. With the constantly increasing rate of cyber-attacks against our nation's network infrastructure and the ever-changing nature of computing, it is vitally important for the DoD to have an understanding of the cyber operating environment in order to properly secure and defend the nation.

More than a decade ago, Bass [76] observed that current intrusion detection technologies were not maturing at the rate of new attacks. Former Director of the National Security Agency (NSA), Mike McConnell, echoed this sentiment in February 2010 when he stated: "The United States is fighting a cyber-war today, and we are losing. It's that simple. As the most wired nation on Earth, we offer the most targets of significance, yet our cyber-defenses are woefully lacking."[77] Commander, United

States Cyber Command (USCYBERCOM) and current Director of the NSA General Keith Alexander continued: "...to defend those networks and make good decision in exercising operational control over them ... will require much greater situational awareness and real-time visibility of intrusions into our networks."[78] These concerns clearly identify the need for a comprehensive strategy to gain situational awareness over the cyber domain, which enables commanders at all levels to consider cyber as they make operational decisions and direct actions for their forces.

To successfully operate in the cyberspace domain, Cyber Situational Awareness (CSA) must be effectively enabled to empower commanders and government leaders to drive action and support rapid decision-making.

In this chapter, I propose six classes of data for establishing situational awareness in cyberspace. Section 3.2 provides background information and motivations for situational awareness. Section 3.3 describes related works in cyberspace research. We describe our data classes in Section 3.4 and present a case study in Section 3.5. Challenges to establishing cyberspace situational awareness are discussed in Section 3.6. Sections 3.7 and 3.8 present conclusions and areas for future research, respectively.

3.2 Background and Motivation

Defining the term 'situational awareness' is almost as hard as actually building situational awareness. United States Department of Defense joint doctrine does not define situational awareness in its Dictionary of Military and Associated Terms, JP 1-02,

though situational awareness is used in the definition of four other terms: blue force tracking, common operational picture, United States Strategic Command's Global Network Operations Center, and national operations center. The closest definition in JP 1-02 was of 'battlespace awareness', but it has been removed from the latest version.

> Battlespace Awareness - Knowledge and understanding of the operational area's environment, factors, and conditions, to include the status of friendly and adversary forces, neutrals and noncombatants, weather and terrain, that enables timely, relevant, comprehensive, and accurate assessments, in order to successfully apply combat power, protect the force, and/or complete the mission.[2]

Since the DoD has established cyberspace as a warfighting domain, many aspects of that definition hold true in cyberspace. With the key being to enable commanders to issue orders to forces based on timely and accurate information. The ultimate goal of situational awareness in cyberspace is to maintain strategic and tactical understanding while continuously taking action or making operational risk decisions.

Achieving CSA has proven difficult to date. However, there are a series of issues to be addressed that will allow incremental progress towards CSA capabilities enabling any organization to harness the power of near real-time information supporting decision-making and proactive actions. Those issues include:

- Identification of what decisions and actions the organization may need to take

with respect to cyber to assure operations can be sustained

- Identification of and access to the appropriate data that supports those decisions and actions
- Analytic tools to make sense of the presented data as it relates to operations
- Technology to consolidate and visualize data for decision makers at multiple levels within the organization

3.3 Related Works

Network defense, and in the military realm, information dominance have been hot topics over the last decade.[79, 80, 81] Computer systems have become fully integrated into our very existence, impacting how we live our lives. Research has been focused on defining cyberspace and developing innovative ways to defend it in the ever-changing cyber environment [82, 83, 84], including discussions focused on the unique challenge that most of the network infrastructure is a commercial product outside the control and protection of any one entity. [83, 85, 86]

There has also been considerable investment into new hardware and software technologies for intrusion detection systems (IDS), host-based security systems, and anti-virus discovery mechanisms. IDS research has moved closer to the individual user and toward a behavioral based approach, as exemplified in [87, 88]. Automated responses have now been included in these detection tools to effectively shut down

an attack once recognized by severing the connection or changing a rule. While progressing, these tools still suffer from a false positive problem which usually causes users to scale back the detection threshold.

Commercial visual analytic tools have been developed in an attempt to provide a CSA picture: IBM's Analyst's Notebook discovers patterns and trends across volumes of data to identify and predict malicious behavior; Palantir's toolset focuses on the fusion of disparate data sources into a unified picture for security analysis; and HP's Arcsite is a security information and event management system for enterprise-level IT architecture. [89, 90, 91, 92] Academic research has also developed visualization techniques in an attempt to provide an insight into the network, most using Ben Shneiderman of the University of Maryland's mantra of "overview first, zoom and filter, and then details-on-demand." [93, 94] VisFlowConnect uses a parallel axes view to the volume of network traffic in sender/receiver pairings over time; CNSSA incorporates information from multiple sources including current vulnerabilities to assign a vulnerability score based on the Common Vulnerability Scoring System; and SiLK provides analysts with the ability to understand, query, and summarize recent and historical network traffic data. [93, 94]

Many publications in the last few years discuss security frameworks to gain insight into the situational environment [84, 95] and even more recently, the notion of tying network security to mission assurance. [83, 96, 97] In [89], the authors present a major task list that a cyber common operating picture must be able to complete as well as

technological concerns in the developing of such a system; the Cyber Attack Modeling and Impact Assessment Framework [98] automates the development of attack graphs for computational analysis and impact assessment; and [99] argues effective policies for near real-time information sharing between multiple parties.

All of these ongoing studies and current analytical tools are inherently important to CSA and the discussion of the optimal way to achieve awareness of the cyber domain; however they do not address the fundamental building block of any situational awareness tool: the data. Our work's novelty springs out of this gap, discussing what classes of information are necessary and how each one builds upon the others to develop a holistic operational picture for establishing situational awareness in cyberspace.

3.4 Cyber Operational Data Classes

To achieve operationally relevant situational awareness of the cyberspace warfighting domain, a system must utilize six classes of information by fusing, correlating, analyzing, and visualizing in near real time. The six classes are as follows: 1) Current and near-future threat environment; 2) Global threats and significant anomalous activity; 3) Vulnerabilities of United States computer systems and underlying infrastructure; 4) Prioritized cyber key terrain that allows understanding of operational and technical risks; 5) Current operational readiness and capability of its cyber forces and sensors; and 6) In-depth knowledge of ongoing operations and critical mission dependencies

on its cyber assets.

As shown in Figure 1, the intersection of any combination of these classes provides more information and moves towards the sweet spot of SA. The factors from all six classes must be continuously assessed in order to provide a true, accurate and holistic representation of the domain which supports the ability to take critical actions and make decisions.

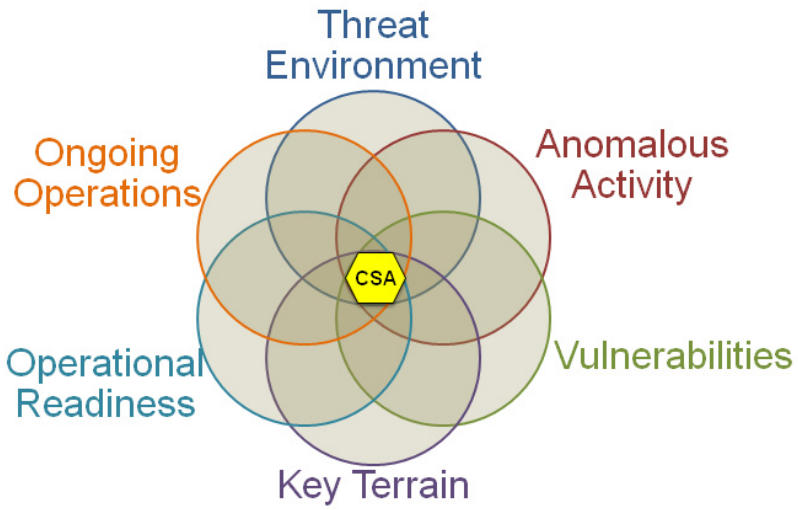


Figure 3.1 : Notional intersection of classes of information requires continuous assessment to provide Cyber SA and enable critical actions and decisions

3.4.1 Threat Environment

To successfully defend the network, an in-depth analysis of potential threats is crucial. This includes an understanding of who would want to attack the network, what goals are they looking to achieve, and how do they normally operate. A thorough knowledge

of a threat's personality and normal behaviors will assist in identifying the threat's tactics, techniques, and procedures (TTP) and developing TTPs for network defense and incident response. Assessing an attack's vector in its early stages may reveal the attacker's capability and behavioral trends, leading to projections of future intrusion activities. This awareness can reap huge rewards in the protection from and reaction to a cyber attack. It also can be used to proactively align resources to counter future attacks using similar TTPs. Development of these adversary profiles could also lead to attribution in the event of an attack.

3.4.2 Anomalous Activity

Most networks have firewalls, anti-virus, and intrusion detection systems, which operate under pre-established rules or signatures, to detect or block when an anomalous activity occurs. These tools cannot respond to a zero-day exploit or a polymorphic virus because these events do not trigger the pre-established rules. Network and host-based IDS are essential to successfully defending the network. However, "IDS sensors can only capture systematic phenomena caused by attacks but cannot positively ascertain whether an attack has happened or succeeded." [79] Baseline historical and current consolidated and normalized data must be incorporated into an automated system in order to understand what is 'normal' and what is 'anomalous' then take actions to effectively defend against cyber threats represented by this activity.

3.4.3 Vulnerabilities

From 2006 to 2011, over 75 thousand new security vulnerabilities were discovered. [100] Vulnerabilities are present in every system no matter how secure the system claims to be. Technology advances so rapidly that it can be virtually impossible to eradicate vulnerabilities altogether. The best one can hope for, in many cases, is simply to minimize them. In order to assess and minimize the risk to the network, vulnerabilities of the systems and the underlying infrastructure must be known. System administrators and security specialists must have the knowledge and tools to understand the vulnerabilities of their networks and to properly test any new system or application before applying it to the network. Most importantly, these vulnerabilities must be known and continuously assessed. Leadership must be willing to allocate funds for vulnerabilities to be found and fixed.

3.4.4 Key Terrain

Though a single organization may have tens of thousands of systems ranging from desktops and mobile devices to routers and switches spread geographically across the world, not all systems have equal criticality to mission success. Defending and garnering full knowledge of all systems, accounts, and processes on the network in real time is impractical. Therefore, it is necessary to identify and prioritize key cyber assets to allow the understanding of critical risks both operationally and technically. Identification of cyber key terrain includes all critical information, systems, and in-

frastructure; whether owned by the organization or used in transit by its information. [101] That said, even these systems must be prioritized and may be less vital than a specific network link supporting a real-time airborne mission. The identification allows for prioritized defense of assets but cannot fail to consider all systems and assets in the network.

3.4.5 Operational Readiness

Organizations must know the operational readiness and capability of their cyber forces and assets. This includes the status of its tools and capabilities along with the ability of its cyber forces to protect its networks. Understanding the training status of all personnel to operate in the current threat environment and the readiness and integrity of network sensors, paths, and systems is critical. A real-time status of the network and personnel resources provides data necessary to recognize an attack and align resources which are available to appropriately respond. Mission impact is another aspect of operational readiness which is often hard to define and keep up to date. For a situational awareness picture to truly be useful, it must be operationally relevant and actionable. For this to occur, an organization must have a thorough understanding of mission dependencies based on cyber assets. With the knowledge and prioritization of intermission and mission-system dependencies, the organization can now depict to leadership the impact of a cyber event, whether an outage or attack, and the significance of securing certain assets. [83, 96]

3.4.6 Ongoing Operations

Lastly, information about the status of all ongoing operations (cyber, kinetic, and even diplomatic) must be fully understood by commanders at all levels. This knowledge could be used to deconflict controlled outages or upgrades to systems that are currently engaged in support of an operation. It could also be used to dynamically identify key terrain and adjust defensive TTPs during the operational window of time. Understanding which operations are being executed or soon to begin execution, allows commanders to reallocate assets as necessary to support those operations. In addition, this allows leaders to understand the operational impact of systems and their critical operational dependencies.

3.5 An Operational Case Study

A hypothetical operational case study is presented in order to emphasize the value of holistic fusion of data from all six classes. In this case study, I introduce a commander and staff whom are initially presented data from the ongoing operations, key terrain, and operational readiness classes. We will show the improved situational awareness opportunities to impact the commander's decision-making process as additional information classes are considered.

A Joint Task Force (JTF) is currently conducting combat operations in an area of operations that requires the continuous flow of logistical and personnel resupply. In the operational planning process, the commander has designated his logistical support

information systems as cyber key terrain. These systems operate on an unclassified military network so they can receive updates from commercial shipping and airflow systems on the Internet. The JTF commander also is aware that the network sensors deployed to protect these logistical systems are degraded due to required maintenance upgrades. The upgrades are currently scheduled for implementation by a computer network defense service provider (CND-SP) stationed in the continental United States during the next month. Lastly, the commander has an extremely proficient cyber investigative and forensics unit attending commercial certification refresher training. With this partial set of information, the commander has a good baseline of situational awareness of cyber assets and how they may impact his operations across all warfighting domains.

During the course of operations, a critical vulnerability in the outdated operating system of the logistical support system is discovered. As a DoD program of record, the potential patch for this vulnerability remains in pre- deployment testing and is not scheduled for release for another 30 days. USCYBERCOM has assessed the vulnerability and issued a high priority message across the DoD cyber enterprise announcing the details of the vulnerability. This vulnerability allows root-level access to be gained on the systems potentially enabling the deployment of malicious software on all unpatched systems. The commander is advised of the potential impact to his key logistics systems, but decides to take no action based on requirements for the continued flow of supplies and personnel supporting his operational mission set.

When the intelligence officer advises the commander on a new cyber threat report, an additional class of data (Threat Environment) is fused with the current understanding of the battlespace. In this report, it is assessed that the adversary has ever-increasing interest in disrupting and influencing the logistical flow of forces and supplies into theater. Additionally, supporting cyber assets are known to deploy Trojan-horse software on susceptible systems. This additional information of the threat environment improves the commander's understanding of the cyber environment and drives him to take decisive action to ensure his combat power will be available at the critical point in his operations. He directs his cyber force to cease with their commercial training and refocus their efforts on monitoring the behaviors of his logistical support platforms.

While reviewing the network flow and log data from the logistical system, the team discovers information included in our last class, Anomalous Activity. More than half of the logistical support systems supporting the JTF have been sending irregular sized traffic over TCP port 443 to a subnet outside of the United States. Further forensics work determines documents have been slowly exfiltrated via covert encrypted and unencrypted channels. The commander is now alarmed and initiates crisis action planning. He directs the stateside CND-SP to immediately upgrade the defensive sensors and remove the logistics systems from the network until appropriate countermeasures can be deployed to protect the systems until the patch becomes available. Further, he requests intelligence and cyber forensics support to determine

which files were stolen and the potential operational impact of their loss. Now that he does not fully trust his logistics systems' information, considering future shipping schedules were the exfiltrated files, he reallocates air and naval assets to protect inbound shipping containers to protect his logistical lines of communications. Lastly, he directs his cyber forces to begin detailed log review with daily update briefings.

This case study portrays an environment where all SA information classes have an abundance of data available for consumption by an integrated system or motivated person able to fuse them together to provide the opportunity for total situational awareness. This is not todays reality. Cyber forces rarely track or concern themselves with the status of ongoing operations across all warfighting domains. Strategic and operational commanders do not know or fully understand how to determine their cyber key terrain. If they do, typically, they have not taken the required actions or time to determine and designate cyber key terrain. Additionally, the operational readiness of cyber forces is not well defined or tracked at the level needed to fully understand capabilities and how it could impact operations. In contrast, vulnerability, threat and anomalous activity data is plentiful within the intelligence and cyber communities. That said, the data is often presented to the commander in a way that information overload or technical jargon routinely make it difficult for the commander to assess the value of the information and therefore the information is discounted or ignored. Other challenges that inhibit today's ability to gain, maintain, and adjust the fusion of information that can provide SA to the commander are described in the

next section.

3.6 Current Challenges

Effective Cyber Situational Awareness requires that data and information be collected, analyzed, and displayed to the end customer in a timely and relevant manner. Although numerous challenges exist, the key barrier to successful implementation and execution of enterprise-wide CSA is solving the following organizational and technical challenges.

3.6.1 Organizational Fear

Gaining access to all of the necessary network data within different aspects of an organization can lead to a turf war. No entity wants to give up access to their data due to fear. Fear of humiliation in publicizing security flaws, fear of losing a competitive edge or public confidence, or fear of the proverbial 1,000 mile hammer. Regardless of the reason, this fear prevents complete situational awareness. To combat this fear, the United States Department of Defense must define and enforce a single information owner who can aggregate this data for analysis.

3.6.2 Data Consolidation & Normalization

Data comes in the form of technical and human collections, including IDS, network sniffers, and computer system log files. Ingesting all of the data is currently impractical but may soon become reality due to the advancement of cloud computing and

the ever increasing data transfer rates. Determining the proper metrics and alert thresholds for the organization are essential for real time analysis. The data from these sources needs to be consolidated and put into a normalized format in order to be properly ingested into a CSA tool. Data refinement is simplified when a common format exists and requires a temporal calibration of the different data streams. [76]

3.6.3 Data Synthesis

Currently, stove-piped data synthesis solutions exist across different parts of organizations that were developed separately over time without a clear coordinated cyber strategy. The challenge arises with how to fuse the data together. The fusion process requires the utilization of processing algorithms, such as Sudit's and Stotz's INFERD system, and comparison with known statistics (from USCERT, MacAfee, Norton, etc) to assess evolving situations and threats in cyberspace. [102] This data synthesis is needed for a full understanding of the normal state of the network, allowing security to move away from signature-based toward true anomaly-based detection. Intruders executing stealth TCP-based attacks on multiple geographically-separated parts of a corporate network may fall below the pre-established security thresholds. A common situational awareness tool which ideally includes all six classes of information may be able to synthesize the data and combine disparate attacks which may paint the picture of a coordinated and sophisticated enemy. [102, 103]

3.6.4 Result Visualization and Dissemination

Until intrusion detection becomes truly machine to machine automation that responds immediately to anomalous activity, human intervention will require rapid understanding by presenting data in a visual manner. In the traditional warfare domains, situational awareness was represented geospatially on a map. Military leadership is used to this representation of disposition of forces, but this depiction does not always fit well within the cyber realm. Visualization systems need to be much more than PowerPoint presentations and bar charts; however, 2D systems such as parallel axes, logical maps, and temporal visualization of packet flows are limited in their ability to represent all the data attributes in one view. In addition, situational awareness visualizations must be able to illustrate mission impact to truly have meaning to leadership. A dissemination plan must also be established for the actionable results as not all information is appropriate for all personnel. Attributes that clearly identify the mission authorities and identity of the user can be used to present the appropriate data to each user.

3.6.5 Timeliness

As the amount of data, rules and signatures increase, analysis accuracy decreases and false positives increase, hampering timely detection and response. Cyber attacks occur frequently and can cause debilitating effects within milliseconds. To combat this, a finely tuned advanced threat detection engine must be used in conjunction

with the known normal state to ensure the broadest possible spectrum of threats are identified and to eliminate false positives as much as possible. The challenge pivots on the ability to summarize vast amounts of information at the appropriate level and then provide it to operators at the appropriate levels in a timely fashion.

3.7 Discussion

The United States' reliance on computer networks is undeniable, and there will never be an impervious defense to all network attacks. Thus, robust situational awareness of the cyber environment, detailing what is happening, where, and what are the best available response options is absolutely critical to operations. In this chapter, I developed a new approach for decision makers to assist in rapid decision making. I introduced six classes of information necessary (threat environment, anomalous activity, vulnerabilities, key terrain, operational readiness and ongoing operations) to effectively enable and empower commanders and government leaders to incorporate cyberspace into the decision making process. This data must be continuously analyzed to provide a true and accurate representation of the domain.

However, there still remain many challenges that must be addressed before situational awareness in cyberspace may be obtained. This chapter has identified the decisions and actions the United States must take with respect to cyber, whether it be analytic tools to correlate the presented data to an operation or the technology to consolidate and visualize data for decision makers. Once addressed, the opera-

tional view of cyberspace can move from one of network assurance to a true mission assurance focused situational awareness picture.

No effective and exhaustive solution exists for recognizing the majority of cyber attacks before they occur and cause damage. With the speed of attack achievable in cyberspace, a fully developed cyber situational awareness picture is as close to an early warning system as one can achieve. Therefore, the challenges must be overcome, and situational awareness in cyberspace must be realized to enable proactive, agile, and successful network defense for the United States.

3.8 Future Work

The classes of data introduced in this chapter are based on my intensive operational experience working in the area of cyber situational awareness for the U.S. Department of Defense. Though I have traveled the world talking about Cyber SA to senior leaders in multiple organizations across the Department, experimentation and prototyping of systems uses these classes is necessary to fully validate the claims.

Several key aspects of attaining situational awareness are still not well defined. Every organization depends on cyber assets to accomplish their mission. These assets can encompass thousands of computer systems, network sensors, and personnel spread across the globe. An efficient method for determining cyber key terrain to assure mission accomplishment has yet to be found.

As networks expand and data rates continue to soar, working with massive datasets

in real time is becoming more common. More research is necessary in taking sensor event data, efficiently storing and correlating it to mission impact, and then disseminating it in a timely manner to enable leadership to make better decisions. The advent of cloud computing may make this more achievable.

Many advances are being made in general data visualization techniques. The conventional SA tool displays network events on a geo-referenced map of the network. This method works well for battlefield awareness in ground, naval, and aerial assets, but may not be the best way to view cyberspace based on interconnections that defy geographic boundaries. Other visualization techniques need to be developed which allow SA at various levels to inform the commanders for leadership decisions and the net defenders or system administrators for decisive actions at the operator or analyst level.

CHAPTER 4

Exploiting Military OPSEC through Open-Source Vulnerabilities

4.1 Introduction

With the ease of social media, military members and their families can post pictures of an ongoing operation or discuss events of a current deployment from any where at any time. While individuals may feel the information publicly posted is harmless when viewed in isolation, it is precisely the type of information an adversary can leverage for use in intelligence targeting. Bits created in one part of the physical world can easily find themselves elsewhere, allowing an adversary to piece together trivial pieces of information to provide a robust picture of activities, friends, and family; all of which can be used, augmented, and aggregated to create a highly refined list of potential intelligence targets. With this era of greater interconnectedness and easier communication, there is a growing tension between military users' personal needs (keeping in touch with family and community) and military OpSec, or operational security (denying an enemy the ability to gather operational intelligence, such as troop movement, as well as making it harder to conduct human intelligence activities at home).

Rather than placing a spy in an organization, adversarial nations or terrorist

groups can use social media to find valuable targets. Since nearly everyone with an Internet connection has a Facebook, Twitter, LinkedIn or other online profile, adversaries can "spy" on members of the United States military; learn their schedules, habits, interests, discontents, secrets, etc; and then bribe, threaten, or coerce them into turning over sensitive information. While this information used to be private, social networking sites provide an adversary an easy avenue for data collection with little risk or cost.

My initial hope was that military members would be well aware of the dangers of social media and, while still using these networks, would ensure appropriate security measures were in place. Although Operational Security (OpSec) is frequently discussed during military training, I was able to recognize many questionable posts on Facebook, LinkedIn, and other social networking sites. While a few non-compliant military members is a concern (as one slip could cost lives), I wanted to determine the amount and type of information available about U.S. military members and how widespread the problem might be.

4.2 Motivation

Based on hundreds of pieces of public and private data ("Big Data"), marketers, financial institutions, and other businesses are creating profiles predicting personal behavior; from how likely we are to buy a product to how likely we are to end up in the hospital [104]. The second largest corporation in the U.S., Google, derives nearly

97% of its revenue by profiling individuals and their search terms and providing relevant on-line advertising [105]. With massive amounts of information freely available or cheaply purchased, what would stop a military adversary of the U.S. from profiling its military members and civilian DoD counterparts to develop intelligence prospects? Of course, similar risks extend to anybody (military or civilian, domestic or foreign). The purpose of this research is to provide quantifiable insights into how the effects of personal social media, mobile, and other Internet use by U.S. military members can result in a high vulnerability of exploitation by an adversary. This study follows Spang [106], which was done largely by hand, where my work uses automated methods to consider a larger sample. My goal was to provide the military with thorough actionable information, possible enemy scenarios, and recommended actions to remediate this real-world threat. I expect this guidance will also be applicable well beyond the U.S. military.

4.3 Department of Defense Guidance

The Department of Defense (DoD) initially attempted to limit the use of open social sites such as Facebook, Flickr, and YouTube due to concerns of security, accountability, and privacy [107]. However, in 2010, the department reversed this ban citing the need for better information sharing and to accommodate younger users who had come to expect social media access, thus allowing use of unclassified .mil computers to access such sites [108]. Whether allowed to use military information systems to ac-

cess these sites or not, social media and social networking have evolved to become the primary communication method used by today's military members and their families, integrated into all aspects of their lives. With this in mind, a review of the guidance provided by the DoD is necessary.

4.3.1 Operational Security (OPSEC)

The Department of Defense Instruction 5205.02E [109] and its associated manual 5205.02M [110] outline the roles and responsibilities of the OpSec program. It defines the OpSec process as a "systematic method used to identify, control, and protect critical information" and defines critical information as "information that the organization has determined is valuable to an adversary" [109, 110]. While not directly addressing social media sites, the instruction calls attention to the exposure of critical information through aggregation and its mitigation through awareness training and guidance for those "using DoD Internet services, other Internet-based capabilities, emerging technologies, or developing information sharing environments that are accessible across the enterprise." In addition, the manual calls for security to be integrated into new systems as well as procedures to deny adversaries the opportunity to take advantage of publicly available information, especially when aggregated.

4.3.2 Chief Information Officers (CIO) Council

The CIO Council[1], established in 2002, is the principal interagency forum to improve practices involving information technology within and between governmental organizations [111]. In 2009, the CIO Council released its "Guidelines for Secure Use of Social Media by Federal Departments and Agencies" [112], followed in 2013 with "Privacy Best Practices for Social Media," [113] advising the federal government on the proper uses and dangers of social media. These two documents emphasize the criticality of cyber security in mission success. They also highlighted that senior management may need ongoing education as barriers are too often perceived as technological rather than communications, strategy, policy, or mismanagement.

Focusing on spear phishing, social engineering, and web application attacks, the CIO Council highlighted a 2009 FBI alert [114] citing social networking sites as a mechanism for attackers to gather information on their targets by harvesting information from publicly accessible sites. They were also concerned with how Facebook apps tend to be unnecessarily granted permissions to read private user information [115]. Their main recommendations for non-official use of social media focused on the need for user training [112, 113].

[1]The CIO Council is a forum to improve agency practices related to the design, acquisition, development, modernization, use, sharing, and performance of federal government information resources. The CIO Council communicates its findings to the Office of Management and Budget and to other executive agencies. [111]

4.3.3 Social Media Handbooks

Following this recognized need for user training, each military service has a social media handbook that derives from DoD Instruction 8550.01 entitled "DoD Internet Services and Internet-based Capabilities" [116]. All of the services see social media as a cheap, effective, and measurable form of communication and even encourage their personnel to use it to share their experiences and to keep it touch when deployed, but personnel must ensure appropriate conduct is maintained between leadership and subordinates. While the majority of each handbook is devoted to the proper methods for official online presences, some tips are provided to military members for unofficial presences.

- Sharing seemingly trivial information online can be dangerous to loved ones and fellow military members, as well as leaving you exposed to identity thieves.
- Do not break OpSec as everything shared online must be considered public.
- Do not geo-tag photos while deployed.
- Differentiate between opinion and official information.
- Be on the lookout for intruders, use appropriate security software, and continually review account and privacy settings.

Notably, Marine Corps provided step by step instructions on securing a Facebook profile; unfortunately, with constant privacy changes made by Facebook, these

instructions were quickly out of date. Each service also emphasized that posting personal information such as children's photos, names, schools, ages, and schedules can be dangerous. Each handbook provided guidance on what types of information not to place on social media; however, these documents did not link social media to the danger of becoming a targeted intelligence asset for an adversary [117, 118, 119, 120].

4.3.4 Guidance of Other Nations

The British and Canadian militaries issue largely similar advice to that given by the United States military, expressing a desire to use social media to help the public understand the challenges the military faces as well as strengthen the bonds between the military member and their family and friends. While they have stressed the need for OpSec within these applications as well as the risk of a simple geo-tagged photo or leaked detail on Twitter, they do not address personal information and how it could be dangerous [121].

The Israeli Defense Forces are much more restrictive. They train their soldiers to not identify themselves or discuss operational issues. They have even instructed their civilians to avoid status updates when rockets or attacks happen nearby, hopefully reducing the possibility of their adversary using social media as a type of battle damage assessment tool [122].

While I was unable to locate Chinese social media guidance, it is important to note that the U.S. (and computer security companies) are profiling Chinese cyber units

through the corroboration of digital attack traces with open source data including social media [123].

4.4 Research Design

This research utilizes select social media sites and uses content analysis and machine learning algorithms to identify members of the United States military (Army, Navy, Marines, and Air Force) who may, based on their Internet activity, be leaking too much sensitive information, becoming a prime target for an adversary's intelligence activities. I began with automated data collection.

4.4.1 Data Collection

Facebook, by its own description, is a social utility that connects people to keep up with friends and share photos and videos [124]. I began there because it is the #1 social media site and the second most visited site in the world based on Alexa's rankings [125]. On Facebook, each user creates a profile detailing as much or as little personal information as they choose, aside from the mandatory fields of name and age. The data initially entered, as well as photos, videos, and other updates provided by the user, can then be shared throughout their network. Depending on how a user chooses to secure their profile, being a member of a user's network offers access to additional information not shared with non-members. The vulnerability I exploited originates from the failure of Facebook users to appropriately choose security

settings preventing the public from viewing potentially personal information [126]. In addition, Facebook's poor default security and constantly changing privacy settings (including many well-publicized fiascos) have left many users with public profiles who may not be aware of exactly how public they are [127].

Using Facebook's application programming interfaces (APIs), I developed a script that searched for users who publicly self-identify as being a member of the United States military and then scraped as much data as possible from their public profile. The script then continued to search, utilizing the identified military member's network of friends, if available. Since the "Employer" category is populated by the user, the field returned varied based on the specific text the user chose to use; for example, members of the U.S. Army may have identified their employer as "US Army", "United States Army", or "U.S. Army". To generate the most comprehensive results possible, I used these and many other variations and merged the results.

All the data I gathered were from public pages. No account intrusions, social engineering, or other deceptive practices were used to gather my data. For better or for worse, Facebook does not verify anything, so users can and will falsely claim to have a military affiliation, present or past. To increase the quality of my results, I manually scrutinized the data and discarded profiles belonging to individuals where inconsistencies were found (i.e., fake sounding name, impossible birth dates, military rank above what is possible for the user's age as age and rank are strongly correlated in the U.S. military, etc). An adversary looking to collect data on U.S. military service

members would likely follow a similar approach. *In all, I found 3080 public Facebook profiles across all four military branches*, with the results of the data discovered shown in Tables 4.1 and 4.2.

. . .

Table 4.1 : Total Number of Facebook Profiles Discovered Broken Out by Military Branch

Facebook						
	Air Force	*Navy*	*Army*	*Marines*	*Total*	*Control*
# of Profiles Found	848	742	1039	451	3080	3386

Next, I applied this same data collection approach to LinkedIn. LinkedIn is a networking site designed primarily to build one's professional identity online, including discovering professional opportunities and business deals, and getting the latest news and insights into a chosen professional industry [128]. It is the third largest social networking site and #12 on Alexa's most visited sites globally, as well as being the self-proclaimed world's largest professional network [125, 128]. Similar to Facebook, the user controls the amount of information entered as well as the amount of information the public or "connections" can view. In addition, LinkedIn offers an upgraded "Recruiter Corporate" account for only $720 per month for full access to all 300 million LinkedIn members [129]. For this research, I only considered LinkedIn data that is visible to the public, without any special permissions or insider access. I note that any intelligence agency would happily spend the additional money for unrestricted access.

Using LinkedIn's APIs, I developed a script that searched for any military mem-

...

Table 4.2 : Percentage of Facebook Profiles Discovered Containing the Specified Information

Facebook						
	Air Force	*Navy*	*Army*	*Marines*	*Total*	*Control*
First Name	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Last Name	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Gender	98.5%	98.1%	99.1%	98.2%	98.6%	98.7%
High School	71.6%	78.0%	84.5%	79.4%	78.6%	84.2%
College/University	64.0%	69.0%	65.8%	54.1%	64.4%	100.0%
Graduate School	12.9%	11.6%	8.0%	8.7%	10.3%	10.7%
Self Photo	72.8%	76.6%	78.6%	82.9%	77.1%	85.8%
Age/DOB	57.9%	56.9%	57.9%	64.5%	58.6%	68.8%
City	53.5%	54.9%	54.8%	50.3%	53.8%	76.7%
Marital Status	32.6%	33.4%	30.4%	29.5%	31.6%	42.9%
Current Place of Work	100.0%	100.0%	100.0%	100.0%	100.0%	45.2%
Spouse Name	11.4%	12.7%	11.4%	13.5%	12.0%	14.0%
Anniversary	1.9%	2.8%	2.4%	1.8%	2.3%	5.0%
Contact Phone	0.2%	0.0%	0.8%	0.7%	0.4%	0.5%
Home Address	0.0%	0.0%	0.1%	0.0%	0.0%	0.2%
Interests	67.8%	68.9%	69.6%	67.2%	68.6%	71.3%
Professional Skills	57.1%	61.7%	54.3%	52.1%	56.5%	51.2%
Viewable Network	79.3%	73.1%	74.7%	74.5%	75.5%	77.8%
Rank	6.1%	7.3%	5.5%	7.1%	6.3%	0.0%
Viewable Timeline	68.8%	71.0%	69.4%	67.4%	69.3%	79.4%
Geotagged Photos	34.4%	43.0%	47.0%	45.0%	42.3%	48.5%

ber's name that had been discovered through Facebook. I then manually corroborated these profiles with the information gained through Facebook to ensure it was the same individual, and any additional information gained was added to the target profile. Any inconsistencies found between the Facebook and LinkedIn data resulted in the exclusion of the target profile altogether. *In all, of the 3080 Facebook profiles I found, 902 had corresponding LinkedIn accounts.* My results are shown in Tables 4.3 and 4.4.

. . .

Table 4.3 : Total Number of LinkedIn Profiles Discovered Broken Out by Military Branch

LinkedIn						
	Air Force	*Navy*	*Army*	*Marines*	*Total*	*Control*
# of Profiles Found	294	248	267	93	902	1217

The results from combining the profiles from Facebook and LinkedIn are shown in Tables 4.5 and 4.6.

4.4.2 Control Group vs. Military

There is no obvious way for us to determine how many military members have social media profiles that are set to less-than-public visibility, or that hide their military affiliation. As a proxy, I decided to consider a "control group" consisting of alumni from two large U.S. public universities. I estimate these two universities to have roughly one million living alumni, yielding a total population of comparable size to the approximately 1.5 million members of the U.S. military. Of course, the age

. . .

Table 4.4 : Percentage of LinkedIn Profiles Discovered Containing the Specified Information

LinkedIn						
	Air Force	*Navy*	*Army*	*Marines*	*Total*	*Control*
First	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Last	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Current City	96.9%	93.2%	95.5%	95.7%	95.3%	95.1%
High School	2.4%	3.6%	5.6%	12.9%	4.8%	9.4%
College/University	70.8%	81.1%	68.2%	64.5%	72.2%	100.0%
Graduate School	34.4%	39.9%	22.9%	21.5%	31.2%	40.8%
Personal Photo	20.1%	19.0%	24.0%	19.4%	20.8%	22.4%
Age	4.4%	6.5%	4.1%	5.4%	5.0%	5.3%
Phone #	0.7%	0.4%	0.0%	2.2%	0.6%	1.2%
E-mail address	6.5%	9.3%	9.7%	8.6%	8.4%	11.0%
Current Place of Work	100.0%	100.0%	100.0%	100.0%	100.0%	93.9%
Work Experience	77.6%	84.3%	83.2%	79.6%	81.3%	87.3%
Professional Skills	49.7%	55.7%	64.8%	51.6%	56.0%	65.2%
Interests	17.7%	23.0%	25.8%	26.9%	22.5%	33.4%
Clearance	17.4%	19.4%	27.0%	20.4%	21.1%	2.0%

. . .

Table 4.5 : Total Number of Facebook and LinkedIn Profiles Discovered Broken Out by Military Branch

Facebook and LinkedIn Combined						
	Air Force	*Navy*	*Army*	*Marines*	*Total*	*Control*
# of Profiles Found	294	248	267	93	902	1217

...

Table 4.6 : Percentage of Combined Facebook and LinkedIn Profiles Discovered Containing the Specified Information

Facebook and LinkedIn Combined						
	Air Force	*Navy*	*Army*	*Marines*	*Total*	*Control*
First	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Last	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Gender	99.0%	98.0%	97.8%	98.9%	98.3%	98.8%
High School	76.5%	77.0%	84.6%	79.6%	79.4%	78.1%
College/University	90.1%	94.4%	87.3%	82.8%	89.7%	100.0%
Graduate School	43.5%	45.6%	26.6%	30.1%	37.7%	41.6%
Personal Photo	77.2%	81.1%	79.8%	86.0%	79.9%	81.1%
Age/DOB	60.2%	59.3%	60.3%	63.4%	60.3%	59.2%
City	99.3%	96.8%	97.4%	96.8%	97.8%	98.4%
Marital Status	34.7%	31.1%	30.3%	23.7%	31.3%	39.9%
Current Placeof Work	100.0%	100.0%	100.0%	100.0%	100.0%	95.7%
Spouse Name	9.5%	12.1%	9.7%	22.6%	11.6%	9.3%
Anniversary	3.1%	4.0%	3.8%	5.4%	3.8%	4.6%
Phone #	0.7%	0.4%	0.8%	3.2%	0.9%	1.3%
Home Address	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%
Interests	73.1%	77.8%	77.5%	68.8%	75.3%	67.6%
Professional Skills	59.9%	62.9%	54.3%	58.1%	58.9%	66.4%
Viewable Network	79.3%	75.8%	74.9%	77.4%	76.8%	79.3%
Rank	4.8%	7.3%	5.2%	7.5%	5.9%	0.1%
Viewable Timeline	64.6%	75.0%	51.3%	48.4%	61.9%	74.9%
Geotagged Photos	34.0%	41.1%	48.7%	47.3%	41.7%	48.2%
E-mail address	6.5%	9.3%	9.7%	8.6%	8.4%	11.0%
Work Experience	77.6%	84.3%	83.2%	79.6%	81.3%	87.3%
Skills	49.7%	55.7%	64.8%	51.6%	56.0%	65.2%
Clearance	17.4%	19.4%	27.0%	20.4%	21.1%	2.0%

demographics will be quite different between these two populations (i.e., as military personnel age they tend to retire from active duty, while alumni never "retire" from their alma mater), so direct comparisons are only meaningful in broad brush strokes.

From my control group, I found 3386 public Facebook profiles, versus the 3080 profiles found from U.S. military members (see Table 4.1). Similarly, from my control group, I found 1217 corresponding public LinkedIn profiles, versus 902 profiles from U.S. military members (see Table 4.5). If I adjust these numbers as percentages of the overall estimated population, I see Facebook and LinkedIn public profiles as 0.21% and 0.06%, respectively, of the military population, and 0.34% and 0.12%, respectively, of the university alumni population.

One can draw a variety of inferences here. Since social network use tends to be more active among younger populations, and the university alumni population skews older than the military population, I can conclude that military personnel are less likely to have public profiles on social networks than the general population. However, these public military profiles seem to share a similar amount of information to the information shared by the university alumni profiles (see, e.g., Table 4.2). From the perspective of military OpSec, it's unacceptable to have *any* such public profiles. An adversary who can convert even one insider can leverage them to a variety of ends. (The recent cases of Bradley Manning and Edward Snowden illustrate how singular individuals may be able to exfiltrate significant volumes of sensitive data.)

Aware of these risks, the U.S. military begins OpSec education in basic training,

with a refresher course required annually thereafter. While this training briefly covers social media sites, it only does so when discussing ongoing operations, not personal risk. With the DoD spending $5.29 million on its operational security program in fiscal year 2014 alone [130], and with the seemingly high number of public profiles of military personnel, I argue that the training program in place is inadequate.

4.4.3 Scoring System

From the information collected above, patterns of life can easily be determined. A good recruiter would have a plethora of this Internet-originated information to draw from to create and foster a fake relationship. With the provided information above, an intelligence officer may easily create a short list of preferred candidates. I used two factors to determine the vulnerability of a United States military member to a foreign intelligence service: 1) the access an adversary has to an individual; 2) the individual's access to classified information.

For each target profile identified in the data collection phase, the following pieces of information, if provided, were recorded as a measure of "Access to Individual" by an adversary: first and last name, schools attended, home address, photos, contact phone, e-mail address, gender, anniversary, rank, professional skills and interests, and viewable network and time line. For "Access to Information", the following pieces of information were recorded and measured: holding a secret or top secret clearance, holding a management position, holding a position allowing for access to information,

and employment with the United States military.

Since not all information is of equal significance to an adversary, I created a weighted scale. I assigned each category a value between 1 and 5 based on the relative importance of the information, with 5 meaning the information would be most valuable to an adversary and 1 meaning the information is least valuable. My weightings are shown in Tables 4.7 and 4.8. This scoring system is a slight adaptation of the one proposed in Spang [106].

The summation of the numeric values for each critical piece of information discovered provided a score for each of the two factors. The combination of the two factors was used to determine the vulnerability of the U.S. military member for exploitation by an adversary (i.e., the higher each factor, the more vulnerable a member was).

Figure 4.1 is a 2-dimensional scatter plot of the numeric scoring results for "Access to Information" vs "Access to Individual" from the combined Facebook and LinkedIn data. The more personal information a member provided contributed to a higher score on the X-axis while more information about a member's access to classified or critical information contributed to a higher score on the Y-axis.

For the "Access to Information" factor, a sufficient baseline is if the member has a security clearance, as represented by the horizontal line at an "access to information" score of 5 in Figure 4.1. This allows the adversary to know the individual has access to restricted areas and classified systems.

On the "Access to Individual" side, the cause for concern has always been if the

Table 4.7 : Numerical-based Scoring System and Rationale for Each Category of Access to Individual Discovered

Access to Individual		
Critical Information	*Assigned Value*	*Rationale*
Home Address	5	Provides direct physical access to individual, probably income
Geotagged Photos	1 - 5	Can provide direct physical access to individuals. 1: less than 5 photos; 2: less than or equal to 15 photos; 3: greater than 15 photos; 4: greater than 15 photos with one location on several occasions; 5: 15 photos and more than 1 location on several occasions
Spouse Name	4	Spouse allows for easy identification; opens another avenue for intelligence; solution to many challenge questions (met them, married them, date married), also allows easier manipulation of primary target
Contact Phone	3	Challenge question answer; provides another avenue for social engineering; Opens potential target of individuals phone
Self Photo	3	Target easily identifiable; may provide insights into interests/family
Viewable Timeline	1 - 3	Can provide information including location, interests, professional skills, friend network, phone number, birthdate, marital status, anniversary, education. Scored on a sliding scale from 1-3 to differentiate between a public wall that's largely empty versus somebody who posts everyday versus somebody who posts personal information (complaining about ex-wife, dislike for job, unhappiness with deployments, etc).
Professional Skills	3	Allows for the narrow targeting of individuals with access to specific data/technology
Age/DOB	2 - 3	Often used for pin numbers, allows for personal identification, often a challenge question. 2 points for age, 3 points for date of birth
Last Name	2	Readily available, easily leads to other data but must be combined with other data to be effective
High school	2	May provide solution to online challenge questions; valuable for social engineering/phishing schemes
College	2	Provides details for social engineer/phishing schemes; also allows determination of possible current job responsibilities
Graduate School	2	Provides details for social engineer/phishing schemes; determination of possible current job responsibilities/level of responsibility
Interests	2	Challenge question answer; also valuable for social engineering/phishing/water hole schemes
Viewable Network	2	Social network allows for larger surface area to find easiest target, as well as providing adversary legitimacy to come after primary target
Anniversary	2	Valuable for social engineering/phishing; also challenge questions
Rank	2	Indication of amount of responsibility, age, salary; can be used during promotion cycles for social engineering/phishing
E-mail address	2	Used as account/user name on many sites, provides avenue for social engineering/phishing
First name	1	Readily available, must be combined with other data to be effective
Gender	1	Easily determinable by first name, not necessarily helpful to enemy
Marital Status	1	Easily found in public records, leads to further investigation of family
Current City	1	Narrows field of potential people to coerce; can be combined with other data to find home address/tax records for specific target

Table 4.8 : Numerical-based Scoring System and Rationale for Each Category of Access to Information Discovered

Access to Information		
Information Category	*Assigned Value*	*Rationale*
Holding a Secret to Top Secret Clearance	5	Clearly states individual has access to classified materials
Holds a position allowing for access to information	2 - 3	Indicates individual probably has access to classified materials. If management position, 3 points, otherwise 2. To determine if management position was held, key terms were searched for in job title, such as commander, director, etc. as well as officer ranks O-3 or higher and enlisted ranks E-7 or higher.
Employed by U.S. Military	1	Indicates individual is military member

person can be identified from the data collected. While none of this data is considered personally identifiable information (PII), technology and the wide availability of information about people enables the aggregation of various pieces of non-PII to produce PII. Sweeney [131] found that the combination of zip code (most of the time easily discerned from city), birth date, and gender was sufficient to uniquely identify 87% of individuals in the United States. With roughly 7.2 billion people in the world, it takes 33 bits of entropy to identify an individual [132]. Gender, for example, is worth one bit; reducing the world's population by roughly half. However, the entropy decreased by other pieces of information is not as easily discernible: living in a heavily populated city or zip code would not be worth as many bits as a sparsely populated city or zip code.

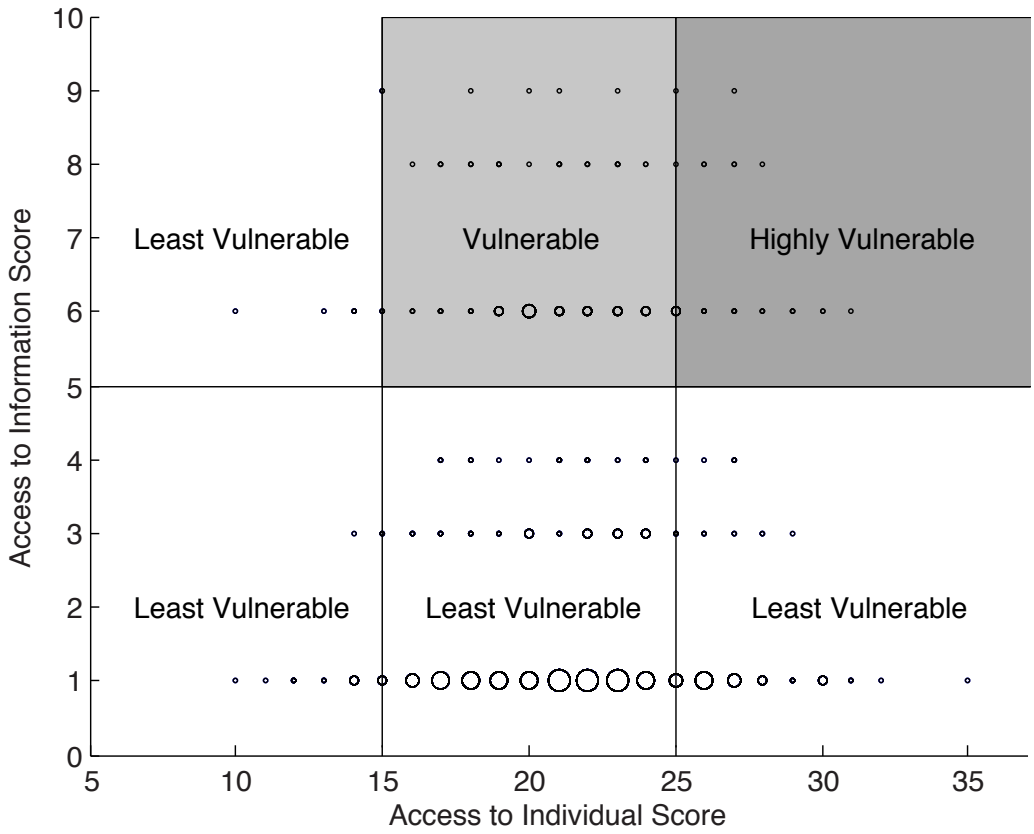


Figure 4.1 : Vulnerability of Military Members via Social Media. Four different sized circles were used corresponding to specific ranges of users within those buckets (larger circle = more users, smaller circle = less users).

Given my weighting scheme, I can declare arbitrary cutoffs above which an individual moves into different categories (I.e., least vulnerable, vulnerable, highly vulnerable). By studying the various criteria and weights that I assigned to them, I considered an "access to information" score about 5 to be a significant threshold. Similarly, I felt that "access to individual" scores above 15 and above 25 seem to be significant differentiators among the people in my dataset. The combination of these

thresholds and the assignment to risk categories is shown in Figure 4.1. With these thresholds, *184 military members are vulnerable and 40 are highly vulnerable.*

I note that absent information from Facebook and LinkedIn might be reconstructible from other sources. Home addresses, for example, can be determined from property records, telephone records, voter registration databases, and a variety of other public sources. Spousal names may be discovered from marriage records. Age, date of birth, and city of birth will appear in birth records. Education records may be public as well (e.g., many unversities publish undergraduate theses online, which will have cover pages indicating author and year; doctoral students produce publications which will also contain contact information). An intelligence analyst would certainly use these other sources to individually evaluate each potential asset. While I did not pursue such a manually-intensive analysis, I wanted to consider how much missing data could be predicted automatically.

4.4.4 Predict Missing Values

Since social media users may omit information about themselves, either out of neglect or concerns for privacy, discovering this additional information makes the profile more of an adversarial intelligence target. Therefore, I wish to apply machine learning algorithms on the information already provided by the user and their network in an attempt to predict the omitted information. To that end, I trained a classifier based on the example of profiles that contained the desired information (label) so it

can be applied to the unlabeled profiles to predict labels for them. To do this, I created a graph representation of the Facebook and LinkedIn social networks collected, $G(V, E, W)$ where:

Nodes V: The set of nodes representing the user profiles collected from Facebook and LinkedIn.

Edges E: An edge $(i, j) \in E$ between two nodes v_i, v_j represents a "friendship" or "connection".

Edge Weights W: The weight w_{ij} on an edge between nodes v_i, v_j indicate the strength of similarities between the two nodes.

Note: Since friendships/connections in Facebook/LinkedIn are reciprocal, the edges in graph G are undirected.

Problem Statement: Given the graph $G(V, E, W)$ with a subset of nodes $V_l \subset V$ labeled and $V_u = V/V_l$ the set of unlabeled nodes (i.e., the information in V not contained in V_l). Let Y be the set of m possible labels, and $Y_l = y_1, y_2, \ldots, y_l$ be the initial labels on nodes in the set V_l. The goal is to infer labels Y on all nodes V of the graph G.

The labels from neighboring nodes are used as the primary source because links between nodes in social networks indicate a relationship between the individuals that the nodes represent. In particular, a link indicates a higher probability of similarity between the linked individuals because friendships usually occur between individuals of similar nature and interests.

Therefore to derive the weights (W) for graph G, the following factors were considered between two nodes i and j: the number of friends i and j have in common, the number of known labels shared by i and j, and how often i and j communicate (i.e., show up in each others' Facebook timeline). Each factor was normalized by its maximum value in G and weighted equally.

To prepare the data for the machine learning algorithms, categorical labels (such as high school, college, graduate school, rank, and current city) were made into individual features using a binary flag for each category. This supervised data set was then used with a variety of standard machine learning algorithms: Naïve Bayes (NB), 3-Nearest Neighbors (3NN), Support Vector Machine (SVM), and Random Forest (RF) [133].

Naïve Bayes (NB) Classification: A standard Naïve Bayes model which for each user-label pair predicts the probability that the label belongs to that user [134].

3-Nearest Neighbors (3NN): A standard k-NN algorithm is used in which the user's label in each category is classified by a majority vote of its three closest (highest weighted) neighbors [135].

Naïve Bayes and 3-Nearest Neighbors algorithms were chosen based on the social phenomena of homophily, which states that individuals tend to associate with individuals similar in nature [136].

Support Vector Machine (SVM): An SVM classification model was built for each label. The Radial Basis Function (RBF) kernel was used and for categorical labels,

a one-versus-many approach was employed. SVMs perform well on data sets with many attributes (labels), even with very few data points on which to train the model. It also has strong regularization properties (less prone to overfitting) which allow it to generalize to new data easily [137].

Random Forest (RF): Random Forest is an ensemble learning method (very closely related to nearest neighbor) for classification that constructs decision trees at training time and outputs the mode of the classes output by the individual trees. Put simplistically, it randomly produces trees of weak learners who work together to form a strong learner. The Random Forest algorithm also does not overfit the data and allows for ranking variable importance in the classification [138].

All my experimentation was performed in a 10-fold cross validation setting, to ensure each model would generalize to an independent data set. To begin each test, the data set of only the profiles where I know the data point were divided into 10 subsamples of equal size, 9 of the subsamples were used as the training set, while 1 was used as the test (or validation) set. The process was repeated 10 times, with each of the 10 subsamples used exactly once as the validation set. In addition, the classification process was made iterative; in other words, the classification process spreads information to new places in the graph which is then fed into the features used for the next round of classification. The results from each of the 10 runs was then averaged to produce a single estimation. For all tasks, test set classification rate (true positive) was reported (as seen in Table 4.9). As an example, the 3NN classifier

had a 89.18% test set prediction accuracy for gender. This means that after being trained, this classifier was able to predict the correct gender of the profiles in the test set 89.18% of the time.

...

Table 4.9 : Test Set Prediction Accuracy for Naïve Bayes, 3-Nearest Neighbor, Support Vector Machine, and Random Forest Classifiers for Each Information Category

	Algorithm			
Information Category	*NB*	*3NN*	*SVM*	*RF*
Age	37.50%	45.96%	34.19%	41.54%
Age (+/- 1)	88.05%	86.58%	81.43%	83.46%
High School	51.26%	50.14%	47.35%	52.37%
College	76.14%	78.99%	74.41%	81.58%
Graduate School	33.82%	43.82%	29.12%	36.18%
Interests	88.51%	92.49%	89.84%	91.90%
Rank	90.57%	88.68%	88.68%	90.57%
Gender	87.94%	89.18%	85.01%	88.50%
Marital Status	78.37%	81.56%	74.11%	77.66%
Current City	65.99%	72.11%	61.68%	69.95%

Intelligence analysis is based on contingent, not absolute, prediction; gauging the likely outcomes based on the range of possible scenarios [139]. From an attacker's perspective, a classifier with 80% or better accuracy is more than sufficient, as perfect information is neither necessary nor ever available. Thus, with these algorithms, additional information such as approximate age, college, interests, rank, gender, and marital status could be inferred for individuals who did not provide all of the details themselves. Combining the profiles from Facebook and LinkedIn with the predictions of the best machine learning algorithms for each category (of those achieving an 80% or better accuracy), the results are shown in Table 4.10.

...

Table 4.10 : Percentage of Profiles Discovered Containing the Specified Information When Facebook, LinkedIn and Best Machine Learning Algorithm Results were Combined

Updated Profiles on Both Facebook and LinkedIn Using Best Machine Learning Classifier			
	Before	*After*	*Difference*
Age (+/-1)	60.31%	95.23%	34.92%
College	89.69%	98.00%	8.31%
Interests	75.28%	98.12%	22.84%
Rank	5.88%	91.02%	85.14%
Gender	98.34%	99.78%	1.44%
Marital Status	31.26%	87.25%	55.99%

As seen in Table 4.9, the classifiers with the best results rely on the person's closest friends in predicting the value of the missing information. It seems, with little surprise, that individuals are closest with friends of the same age, background, interests, rank, and marital status (i.e. Academy graduates tend to still be close friends with other Academy graduates and married individuals tend to associate more with married individuals). *With the additional information from the machine learning classifiers, 57 more members moved into the highly vulnerable category, 4 into the vulnerable category.* If I had a ground truth to create a classifier for the clearance category, my results would have been much higher as approximately 79% of the profiles collected were immediately deemed not vulnerable due to this constraint.

4.4.5 Finding Other Military Members

For the final aspect of this research, my goal was to increase the number of possible targets found through social media. Again, machine learning algorithms were used in an attempt to recognize individuals who do not self-identify as a member of the military but who actually are. To do this, the same graph setup as before was used, however, this time it was limited to members of the U.S. Air Force and their associated connections; as I were able to verify current employment through an Air Force internal human resources database. A classifier was trained based on the example of profiles that contained current employment so it can be applied to the unlabeled profiles in order to predict whether or not they were currently employed by the Air Force. Using only the publicly available social network connections from the 294 self-identified Air Force employees who had both a Facebook and LinkedIn profile, 87 new targets were discovered. Of these, 74 turned out to be current Air Force members. (For this study, the internal Air Force database was not used for any other purpose beyond testing false-positive rate.) Overall, this study achieved an accuracy of 85.3% in predicting military status with a false positive rate of 14.7%. Scoring these profiles with the system described in Tables 4.7 and 4.8, I determined 13 of the 74 were vulnerable, 4 highly vulnerable.

If these same ratios held for the other three military branches, this search of publicly available social profiles on Facebook and LinkedIn, combined with machine learning algorithms, would have *resulted in a total of 1168 potential intelligence tar-*

gets, of which 223 are vulnerable, with 109 highly vulnerable.

4.5 Scenarios

A few hypothetical scenarios are presented in order to emphasize the negative consequences that could arise from the massive amounts of data easily gathered about our service members.

Tagging Military Targets While Traveling: Intelligence adversaries use this database gleamed from social media to immediately identify a U.S. military member, while in civilian clothes, on vacation in a foreign country (using additional information from hotel reservation, credit cards, passport, etc). The member is then captured and held for intelligence or ransom. This may be exceptionally dangerous as military members frequently travel overseas on civilian carriers.

Faux Employment: A military officer has complained multiple times on Facebook account about financial problems and has begun looking for outside work income using a LinkedIn account. Knowing the officer's background and access, an adversary uses a consulting firm as a guise to engage with the officer and gather information on U.S. military capabilities and tactics.

Prisoner of War / Kidnapping: An officer is captured while in active duty combat. According to Article 5 of Military Code of Conduct, the officer should only disclose name, rank, serial number, and date of birth to the captors. However, through social media, the captors have access to family names and pictures as well as social

network commentary (e.g., about the current war, or about broader political opinions), allowing them to manipulate their captive as well as to broadcast more effective propaganda.

Watering Hole/Phishing Scheme for Cyber Intrusion: With cyber espionage being a convenient and powerful option for many adversaries, they will naturally employ it to gain insight into U.S. capabilities. Through social media, they may infer a soldier's rank, base, and operational assignment. Knowing that sodier's friends, school history, and current interests, they are fully armed to mount an effective and targeted email attack (i.e., spear phishing) or even an in-person attack (i.e., watering hole) against the officer. Kaspersky Labs assesses this to be a prominent attack vector. Over 20% of the 37.3 million phishing attacks in 2012-2013 occurred through social media [140].

4.6 Recommended Actions

Social media enables immense capabilities for the military through easy dissemination of information and increasing esprit-de-corps. This paper does not refute this nor does it recommend going back to a world without social media, as personal and mobile use of these social platforms is immense; banning their use would not necessarily solve these problems. However, certain risk mitigation strategies must be in place to protect the personal data of our military members as well as the security of our military operations. I present a range of options, recognizing that not all of them will

be palatable.

As it stands right now, users are the weakest link, as they are inadvertently divulging personal (and possibly sensitive) information through social media. Few technical controls can defend against clever social engineering attacks, whether phishing, faux profiles, or watering holes. Therefore, the threat caused by social media must be made aware to military members through periodic training of policy, guidance, and best practices; a start would be its inclusion, or stronger emphasis, in the annual information assurance and OpSec training. This training should include use cases of the dangers posed. Members should also be advised to secure and possibly scrub their profiles prior to deployments.

Another option would be integrating social media into the security clearance investigative process (notably, there are no social media questions in the current paperwork). It's possible that even very basic questions (e.g., "do you have a Facebook profile? Is it visible to the public?"), perhaps integrated with the efforts of the agents conducting the clearance investigations, would help military users of these services to realize when they've inadvertently made their public profiles too revealing.

Furthermore, an obvious approach would be for the military to continuously conduct "opposition research" on itself through the methodology described in this paper. A continuous process like this would mean that "friendlies" would hopefully discover these vulnerabilities and move to correct them prior to adversaries exercising their own opportunities. This might also be combined with "red team" exercises in a va-

riety of forms (e.g., a red team might use social media as part of a social engineering effort to gain unauthorized access).

There's also a role for regulatory action, which might restrict the "default" public visibility of user profiles, or might constrain the ability for third-party brokers to buy, sell, and aggregate such data without user consent. While such data aggregation services might be primarily of use to Internet advertisers, intelligence agencies might also be able to benefit from the low cost data flows in these environments. (A full study of the degree to which cookies and other advertising identifiers can be used to develop profiles on military users would be an excellent area for future research.)

4.7 Related Work

I now survey other related efforts to study the privacy risks of personal social network usage by military personnel.

4.7.1 Military Research

Between 2006 and 2007, the U.S. Army Web Risk Assessment Cell studied 2.5 million official web pages and 2 million private pages, challenging the traditional assumption that soldiers' personal blogs and websites posed a greater OpSec threat than official Army sites. In all, the cell found over 1100 OpSec violations, with 1095 violations on official Army sites. Examples of violations included schematics of communications networks, including names of servers and routers; classified documents; detailed maps

of training areas and facilities; and personally identifiable information on individual soldiers and their families [141]. While this study quantified OpSec violations in official Army websites and highlighted the need for further controls or increased training, it occurred in the infancy of social and mobile media.

In 2007, an Air Force study looked at 500 MySpace profiles belonging to Air Force members. The audit tallied up the total amount of information provided, with each piece of information (first name, last name, home town, etc) being equally valued. If a user provided more than 6 pieces of information, these users were deemed good targets, with 60% of profiles visited falling into this category [142]. While limited in scope and size, the study quantified the percent of Air Force members directly vulnerable to targeting. However, it did not address the individual's access to further information and the vulnerability inherent in that subsequent access.

4.7.2 Joseph Spang's Thesis

As the precursor and starting point for this research, Spang looked at each agency in the Intelligence Community and manually searched Facebook and LinkedIn for up to 50 profiles per agency. The individual did not have to possess both a Facebook and a LinkedIn account to be used and due to fraudulent[2] profiles, some agencies had less than 50 valid profiles. Therefore, a total of 658 profiles were found. Ultimately, 83 profiles or 13% were deemed vulnerable since they provided enough information

[2]In this study, profiles were deemed fraudulent if the totality of the information provided suggested the users were most likely not employees of given agency (e.g., they did not live anywhere near agency claimed to work for).

to determine access to classified information and enough personal information to be fully identified with additional open source research. Although a small study, the results showed the practicality of leveraging social networking sites to identify vulnerable employees of the intelligence community, cutting down the amount of time and resources an enemy would need to allocate in their intelligence cycle [106].

4.7.3 Data Leakage on Internet-based Platforms

Data Brokers

Data brokers, companies that specialize in gathering information about consumers and selling it to a third party, have turned the development of consumer profiles into a multi-billion dollar industry. There are hundreds of these companies, analyzing everything from our online search history, to our brick-and-mortar store purchases, as well as our income and debt levels, and religious and political affiliations [143].

Retailers sell transactional data to marketers looking to target specific areas for ads. If you have ever signed up for a loyalty card, entered a sweepstakes or filled out a survey, all of the information is probably up for sale. Datalogix, for instance, collects information from store loyalty cards for more than $1 trillion in consumer spending across 1400+ leading brands [144]. These companies then take this information to sort people into groups and then sell it to marketers, employers, charities, governments, and other businesses. Another broker, Spokeo.com, will for a small fee provide an individual's home address and phone, mobile phone, lifestyle and interests, email

address, and even financial information, simply by searching a name, email address, or phone number [106].

Other than certain kinds of protected data, including medical records and credit reports, U.S. consumers have no legal right to control or even monitor how information about them is being used. Federal Trade Commission Chairwoman Edith Ramirez stated "The extent of consumer profiling today means that data brokers often know as much–or even more--about us than our family and friends" [145]. While most of this data is used to sell products, other information is bought to screen prospective employees, determine our likelihood to wind up in the hospital, or our predicted loyalty to our current cable company. In the 2012 elections, companies linked voting records with computer cookies, allowing candidates to target online ads based on whether the user was a registered Democrat or Republican or how much they donated to political campaigns before [143]. Additionally, some data brokers have partnered with social media sites to collect "screen names, website addresses, interests, hometown and professional history, and how many friends or follows you have" [143].

Mobile Data Leakage

The U.S. military is assessing the potential of mobile devices for use in missions ranging from marking locations of roadside bombs, sharing intelligence, and even flying drones [146, 147]. With this in mind, and knowing that 58% of adults in the United States use a smart phone, it is essential to also look at data leakage through

these mobile devices [148].

As of October 2013, there were over one million Android-based malicious applications, most disguised as legitimate applications. These applications often request unnecessary permissions (e.g., location, communication, accounts access, enumerating the other apps installed on the device,, and phone call records). In addition, they may pose as a legitimate financial application, siphoning off a user's banking data [149]. Similar issues occur on Apple iOS devices, including malicious apps stealing contact information [150] as well as secretly logging user touch inputs and login credentials [151].

Even with legitimate applications, there is a substantial risk of personal data leakage. In Android devices, ad libraries are provided the same privileges granted to the application. Thus, advertisers are provided the ability to obtain a user's location, contact list, and account names, take pictures or video with their camera, or even record their conversations, so long as the host application has those permissions. It is important to note that the ability of an ad library to use these permissions does not mean they are being abused. However given an application developer's incentive to maximize ad revenue, they have a corresponding incentive to leak private user data [152, 153].

With all of this data, data brokers can determine people experiencing life-changing events such as getting married, buying a home, sending a kid to college, or getting divorced—opening up insight into when people are at a moment of possible financial

or emotional hardship in their lives [143]. Former White House Chief Information Officer Theresa Payton stated "I truly believe government agencies and businesses don't want to pry on our personal lives, but all databases are hackable. This data is a gold mine for cyber criminals" [1]. And all of the data available to a cyber criminal is equally available to a military adversary.

4.7.4 Counterintelligence

Counterintelligence (CI) refers to "efforts taken to protect one's own intelligence operations from penetration and disruption by hostile nations or their intelligence services" [33]. Most nations have an intelligence apparatus of some sort because knowing what the other side knows, does not know, and how they operate, can be very useful. The intelligence community and the military services that operate in that community, therefore, become a target for other nations as they attempt to collect, process, and exploit any information to their advantage, possibly using this information to enlist or strong-arm spies.

While the U.S. military has established a series of internal processes to weed out potential applicants or current employees whose loyalty and past activities are questionable, in today's globally connected world, does it go far enough? The background investigation for a clearance request examines a person's last 7 years of activities including personal financial and medical records and possibly interviewing family and friends. This investigation assumes that previous illegal activity was discov-

ered and documented or that ill-gotten financial gains show up in some way that is detectable. Other than the initial investigation, the military uses information compartmentalization--allowing an individual access only to information necessary to perform his/her duties, not having access to all intelligence information available. While "need to know" has been the standard, in the aftermath of the September 11 attacks, then Director of National Intelligence Mike McConnell changed this standard by emphasizing a "responsibility to provide" standard, to ensure sharing is occurring within and across governmental agencies in 2007 [33].

4.8 Discussion

The Internet was designed with the concepts of reliability and free flow of data, linking all corners of the globe together using common protocols. In wiring together the planet, it has also provided U.S. adversaries a fast, adaptable operating environment that provides a significant cost-benefit advantage relative to traditional targeting techniques. I have shown how a foreign intelligence service or non-state actor can easily discover United States military members through simple open-source querying of social media. While my methodology was limited to passive observation of publicly available information and machine learning algorithms, it produced a rich target set.

Through open source Facebook and LinkedIn data, I found 184 vulnerable military members, with 40 classified as highly vulnerable. Further, I expanded these numbers by inferring omitted profile details through machine learning techniques, resulting

in a total of 1168 potential intelligence targets, of which 223 are vulnerable, with 109 highly vulnerable. Further, unstructured open-source research or merging the results with information purchased from a data broker could have further boosted these results. I find, consistent with Spang [106], that, *military members provide the type of personal and professional information in their social network that adversaries spend years and significant resources attempting to develop.*

It is important to note that the portion of the digital corpus of information holding potential analytic value is growing. According to an International Data Corporation study in 2012, 25% of the digital universe contained information that might be valuable if analyzed, with only 0.5% currently being analyzed [154]. The marketing gold rush for Big Data will ensure untapped data will be discovered. Data are being collected about when individuals are home, their heart rate during a run, or how well they are sleeping at night. When fully realized, the full implications of accumulating health, browsing history, purchasing habits, social behaviors, religious and political affiliations, and finances are not well understood. This data accumulation will have important privacy ramifications for civilians, and even more so for military personnel. This makes it essential for the military to track the state of the art in this space and stay ahead of its potential use to target military personnel.

This chapter focused on the availability of open source data in identifying and prioritizing members of the United States Armed Forces from a counter-intelligence perspective. However, recent threats from terrorist organizations bring to light an-

other disturbing revelation of this data's potential use. A July 2014 Law Enforcement Bulleting warned of a "continued call - by Western fighters in Syria and terrorist organizations - for lone offender attacks against U.S. military facilities and personnel" followed by a call by Islamic State militants to "scour social media for addresses of [military] family members and to show up and slaughter them" [155]. U.S. military personnel must remain vigilant and work to better secure their privacy.

4.9 Disclaimer

The research topic, as well as all thoughts, ideas, and recommendations in this chapter were those of the authors, not of the United States Department of Defense, Rice University, or the University of Houston. All data, though publicly accessible, was stored offline and encrypted using AES-256 for protection once aggregated. Once the research was completed, the external hard drive was wiped and physically destroyed.

CHAPTER 5

Conclusion

The addition of cyber into the realm of warfare has blurred the edge of the battlefield bringing about new challenges to traditional well-understood military concepts. In this thesis, I have discussed a major concern at each level of warfare, addressed the challenges involved as well as presented possible solutions.

At the strategic level, the instantiation of the Cyber Domain in 2011 declared cyberspace as having equal standing alongside land, sea, air, and space as a domain of warfare. However, in doing so, no clear guidance has come forth explaining how cyber fits into the traditional military structure or any doctrine concerning cyber's role in future conflicts. Cyberspace presents our adversaries with a cheap and effective means avenue to overcome the tremendous advantage the United States currently enjoys in traditional warfare. Since cyber is going to be a part of any future conflict, if we do not realize the unique aspects of the cyber domain and define a policy and doctrine around it that fully integrates these unique aspects into the military structure, we are in trouble.

At the operational level, I analyzed current situational awareness tools and discussed with current military commanders about their needs in understanding the cyber battlespace. With these lessons learned, a holistic operational framework con-

sisting of six classes of information was developed that once fused, correlated, analyzed and visualized in real time would provide commanders the ability to maintain strategic and tactical understanding in cyberspace. The six classes are as follows:

1. Current and near-future threat environment;
2. Identify global threats and significant anomalous activity;
3. Vulnerabilities of our nations computer systems and underlying infrastructure;
4. Prioritized cyber key terrain that allows understanding of operational and technical risks;
5. Current operational readiness and capability of our cyber forces and sensors
6. In-depth knowledge of ongoing operations and critical mission dependencies on our cyber assets.

As the military has cyberspace so integrated into its daily operations, Situational Awareness in Cyberspace is absolutely critical to not only cyber operations but operations in the traditional domains of land, sea, air and space.

At the tactical level, the rise of social media has created a growing tension between military users' personal needs and military operational security. Adversaries of the United States are using these sites to piece together seemingly trivial bits of information to pinpoint potential intelligence targets. In this thesis, I designed an automated approach to determine the amount of openly available information provided

by U.S. military members through social media; analyzed it through content analysis; then applied machine learning techniques to learn as much from the provided data as possible. In doing so, I determined the current state of DoD policy regarding social media is lacking. I discovered over 1,100 potentially intelligence targets, hypothesized about potential negative scenarios in which this data could be used; and provided recommended actions. With enemy states as well as terrorist organizations using this publicly available resource, without proper attention from the military operational security standpoint, our military members are at risk.

The United States relies on networked computing for all manner of economic, social, civic activity. As such, the military has integrated these networks into every aspect of how it protects the nation and conducts warfare. However, casually applying well-known concepts from physical space without understanding the unique aspects of cyberspace, is a recipe for failure. The United States Military must address these aspects and define a policy and doctrine around it that fully integrates these unique aspects into the military structure.

BIBLIOGRAPHY

[1] M. Dunn, "Levels of war: Just a set of labels?" Research and Analysis: Newsletter of the Directorate of Army Research and Analysis, Oct 1996.

[2] Department of Defense, *Joint Publication 1-02 Dictionary of Military and Associated Terms*, 2010. [Online]. Available: http://www.dtic.mil/doctrine/new_pubs/jp1_02.pdf

[3] J. Arquilla and D. Ronfeldt, "Cyberwar is coming!" *Comparative Strategy*, 1993.

[4] T. Rid, "Cyber war will not take place," *Journal of Strategic Studies*, vol. 35, no. 1, pp. 5–32, 2012.

[5] D. Albright, P. Brennan, and C. Walrond, *Did Stuxnet Take Out 1000 Centrifuges at the Natanz Enrichment Plan? Preliminary Assessment*, 2010.

[6] I. Thompson, "Snowden: US and Israel did create stuxnet attack code," *The Register*, 2013.

[7] M. V. Hayden, "The future of things 'cyber'," *Strategic Studies Quarterly*, vol. 5, no. 1, pp. 3–7, 2011.

[8] Department of Defense, *Joint Publication 3-12 Cyberspace Operations*, Department of Defense, Feb 2013. [Online]. Available: http://www.dtic.mil/doctrine/new_pubs/jp3_12R.pdf

[9] N. Wiener, *Cybernetics, or Control and Communications in the Animal and Machine*. MIT Press, 1948.

[10] W. Gibson, *Neuromancer*. Ace Books, 1984.

[11] G. Shteyngart, *Super Sad True Love Story*. Random House, 2010.

[12] S. C. Butler, "Refocusing cyber warfare thought," *Air and Space Power Journal*, 2013.

[13] Department of Defense, *Cyberspace Policy Report: A Report to Congress Pursuant to the National Defense Authorization Act for Fiscal Year 2011, Section 934*, 2011. [Online]. Available: http://www.defense.gov/home/features/2011/0411_cyberstrategy/docs/

[14] (2013, Nov) The american heritage dictionary of the english language. Houghton Mifflin Harcourt. [Online]. Available: http://www.ahdictionary.com

[15] J. Sheldon, "Deciphering cyberpower: Strategic purpose in peace and war," *Strategic Studies Quarterly*, 2011.

[16] R. R. Raines, *Getting the Message Through: A Branch History of the U.S. Army Signal Corps*, 1996. [Online]. Available: http://www.history.army.mil/html/books/030/30-17-1/CMH_Pub_30-17-1.pdf

[17] Department of Defense, *Joint Publication 3-0 Joint Operations*, 2001. [Online]. Available: http://www.fs.fed.us/fire/doctrine/genesis-and-evolution/source-materials/dod-joint-ops-doctrine.pdf

[18] M. C. Libicki, *Conquest in Cyberspace: National Security and Information Warfare*. Cambridge University Press, 2007.

[19] Department of Defense, *Joint Publication 3-0 Joint Operations*, 2006. [Online]. Available: http://www.dtic.mil/doctrine/docnet/courses/operations/jfcon/jp3_0.pdf

[20] ——, *Strategy for Operating in Cyberspace*, 2011. [Online]. Available: http://www.defense.gov/news/d20110714cyber.pdf

[21] M. Libicki, "Why cyber war will not and should not have its grand strategist," *Strategic Studies Quarterly*, 2014.

[22] R. A. Clarke and R. Knake, *Cyber War*. HarperCollins, 2010.

[23] B. Schneier, "There's no real difference between online espionage and online attack," *The Atlantic*, 2014.

[24] J. Nye, *The Future of Power*. Public Affairs, 2011.

[25] M. Stytz and S. Banks, "Toward attaining cyber dominance," *Strategic Studies Quarterly*, 2014.

[26] O. A. Hathaway, R. Crootof, P. Levitz, H. Nix, A. Nowlan, W. Perdue, and J. Spiegel, *The Law of Cyber-Attack*, Yale University, 2012. [Online]. Available: http://www.law.yale.edu/documents/pdf/cglc/LawOfCyberAttack.pdf

[27] J. Arquilla and D. Ronfeldt, "Cyberwar is already upon us," *Foreign Policy*, Feb 2012. [Online]. Available: http://www.foreignpolicy.com/articles/2012/02/27/cyberwar_is_already_upon_us

[28] M. Boot, *War Made New: Technology, Warfare, and the Course of History 1500 to Today*. Gotham Books, 2006.

[29] B. Donohue. (2013, Feb) How much does a botnet cost? ThreatPost. [Online]. Available: www.threatpost.com/how-much-does-botnet-cost-02

[30] A. Greensberg, "Shopping for zero-days: A price list for hackers' secret software exploits," *Forbes*, 2012. [Online]. Available: http://www.forbes.com/sites/andygreenberg/2012/03/23/shopping-for-zero-days-an-price-list-for-hackers-secret-software-exploits/

[31] J. Menn, "Booming 'zero-day' trade has Washington cyber experts worried," *Reuters*, 2013. [Online]. Available: www.reuters.com

[32] C. Purefoy. (2012, Jan) Underwater cables bring faster internet to west africa. CNN. [Online]. Available: www.cnn.com

[33] M. M. Lowenthal, *Intelligence: From Secrets to Policy*. CQ Press, 2009.

[34] P. Piper, "Nets of terror: Terrorist activity on the internet," *Searcher*, 2008.

[35] J. Pace. (2013, June) Obama presses China's leadership on cybersecurity issues. Associated Press. [Online]. Available: www.nydailynews.com

[36] W. J. Lynn, *Cyber Security - Defending a New Domain*, 2010. [Online]. Available: http://www.defense.gov/home/features/2010/0410_cybersec/lynn-article

[37] PC Magazine. (2014) Dod cyberspace glossary. [Online]. Available: http://www.pcmag.com/encyclopedia/term/62535/dod-cyberspace-clossary

[38] J. Keegan, *Intelligence in War*. Alfred A Knopf, 2003.

[39] D. Kahn, "An historical theory of intelligence," *Intelligence and National Security*, vol. 16, pp. 79–92, 2001.

[40] C. Guglielmo. (2013, Feb) Cisco mobile data shows surge in smartphone users, 4g usage. Forbes. [Online]. Available: www.forbes.com/sites/connieguglielmo/2013/02

[41] M. C. Libicki, *Cyberdeterence and Cyberwar*. RAND Corporation, 2009.

[42] D. Brown, "Resilient botnet command and control with Tor," *DefCon*, 2010.

[43] C. Crowley, "Tor-nonymous - using Tor for pen testing," *Blog: Sans Penetration Testing*, 2014. [Online]. Available: www.pentesting.sans.org/blog/pen-testing

[44] E. Bumiller and T. Shanker, "Panetta warns of dire threat of cyberattack on US," *The New York Times*, 2012. [Online]. Available: http://www.nytimes.com/2012/10/12/world/panetta-warns-of-dire-threat-of-cyberattack.html?pagewanted=all

[45] S. Weinberger. (2014, Oct) How israel spoofed syria's air defense system. Wired.com. [Online]. Available: http://www.wired.com/dangerroom/2007/10/how-israel-spoo/

[46] S. Farole, "The limits of military power: Why the Pentagon's cyber roes can't (and shouldn't) solve America's cybersecurity problems," *Center for Strategic and International Studies*, 2012.

[47] S. Liles, M. Rogers, E. J. Dietz, and D. Larson, "Applying traditional military principles to cyber warfare," in *4th International Conference on Cyber Conflict*, 2012.

[48] E. Sterner, "Retaliatory deterrence in cyberspace," *Strategic Studies Quarterly*, 2011.

[49] T. Franz, "The cyber warfare professional: Realizations for developing the next generation," *Air and Space Power Journal*, 2011.

[50] G. Conti and D. Raymond, "Leadership of cyber warriors: Enduring principles and new directions," *Small Wars Journal*, 2011.

[51] M. C. Libicki, "Cyberspace is not a warfighting domain," *I/S: A Journal of Law and Policy for the Information Society*, 2012.

[52] B. Krebs. (2014, January) System security: In depth security news and investigation. [Online]. Available: www.krebsonsecurity.com

[53] Department of Defense, *Memorandum of Agreement Between the Department of Homeland Security and the Department of Defense Regarding Cybersecurity*, 2010. [Online]. Available: https://www.dhs.gov/xlibrary/assets/20101013-dod-dhs-cyber-moa.pdf

[54] H. S. Lin, "Defining self-defense for the private sector in cyberspace," *World Politics Review*, 2013.

[55] K. Dilanian, "U.s. chamber of commerce leads defeat of cyber security bill," *Los Angeles Times*, 2012. [Online]. Available: http://articles.latimes.com/2012/aug/03/nation/la-na-cyber-security-20120803

[56] J. Garamone, "Panetta spells out DoD rules in cyberdefense," *Armed Forces Press*, 2012.

[57] Villanova Law, "Law review," 2012. [Online]. Available: http://lawweb2009.law.villanova.edu/lawreview/wp-content/uploads/2012/01/VLR315.pdf

[58] Executive Office of the President, *International Strategy for Cyberspace: Prosperity, Security, and Openness in a Networked World*, 2011.

[59] K. Alexander, "Warfighting in cyberspace," *Joint Forces Quarterly*, 2007.

[60] United Nations, "General assembly resolution 3314," 1974. [Online]. Available: https://www1.umn.edu/humanrts/instree/GAres3314.html

[61] United Nations Office for Disarmament Affairs, *Developments in the Field of Information and Telecommunications in the Context of International Security*, United Nations, Jun 2013. [Online]. Available: http://www.un.org/disarmament/HomePage/ODAPublications/DisarmamentStudySeries/P

[62] Council of Europe. (2001, November) Convention on cybercrime. [Online]. Available: http://conventions.coe.int/Treaty/Commun/QueVoulezVous.asp?CL=ENG&NT=185

[63] R. Hurwitz, *An Augmented Summary of The Harvard, MIT and U. of Toronto Cyber Norms Workshop: October 19-21, 2011*, 2012.

[64] J. Rabkin and A. Rabkin, "For new challenges, revisit old rules: Cyber aattack and the law of armed conflict," 2014. [Online]. Available: www.cs.princeton.edu

[65] M. N. Schmitt, *Tallinn Manual on the International Law Applicable to Cyber Warfare*. Cambridge University Press, 2013.

[66] Wikipedia. (2014, Jan) Iraqi air force. [Online]. Available: www.wikipedia.com/wiki/Iraqi_Air_Force

[67] J. A. Ophardt, "Cyber warfare and the crime of aggression: The need for individual accountability on tomorrow's battlefield," *Duke Law and Technology Review*, vol. 3, 2010.

[68] J. Horgan, "What ancient Greeks can teach us about drones and cyber-war," *Scientific American*, 2012. [Online]. Available: http://blogs.scientificamerican.com/cross-check/2012/06/12/what-ancient-greeks-can-teach-us-about-drones-and-cyber-war/

[69] (2014, Nov) Computer virus families: Origins and differences. [Online]. Available: http://www.net-security.org/virus_news.php?id=127

[70] R. Clarke and S. Andreasen, "Cyberwar's threat does not justify a new policy of nuclear deterrence," *The Washington Post*, 2013. [Online]. Available: http://www.washingtonpost.com/opinions/cyberwars-threat-does-not-justify-a-new-policy-of-nuclear-deterrence/2013/06/14/91c01bb6-d50e-11e2-a73e-826d299ff459_story.html

[71] A. W. Vacca, "Military culture and cyber security," *Survival: Global Politics and Strategy*, 2011.

[72] C. Woods and A. Ross, "Drone warfare revealed: Us and britain launched 1200 drone strikes in recent wars," *The Bureau of Investigative Journalism*, 2012.

[73] United States Government Accountability Office, *Cybersecurity: A Better Defined and Implemented National Strategy is Needed to Address Persistent Challenges*, 2013.

[74] Air Force Historical Research Agency. (2014, Jan) The birth of the United States Air Force. [Online]. Available: http://www.afhra.af.mil/facsheets/facsheet.asp?id=10944

[75] (2014, Nov) Ray bradbury online. [Online]. Available: http://www.spaceagecity.com/bradbury/quotes.htm

[76] T. Bass, "Intrusion detection systems & multisensor data fusion: Creating cyberspace situational awareness," *Communications of the ACM*, 2000.

[77] M. McConnell, "Mike mcconnell on how to win the cyber-war we're losing," *Washington Post*, 2010. [Online]. Available: http://www.washingtonpost.com/wp-dyn/content/article/2010/02/25/AR2010022502493.html

[78] K. Alexander, "Advance questions for Lieutenant General Keith Alexander, USA nominee for Commander, United States Cyber Command," *Washington Post*, 2010. [Online]. Available: http://www.washingtonpost.com/wp-srv/politics/documents/questions.pdf

[79] J. Li, Z. Ou, and R. Rajagopaian, "Uncertainty and risk management in cyber situational awareness," in *Cyber Situational Awareness*, 2010.

[80] C. Croom, "The defenders 'kill chain'," *Military Information Technology*, 2010.

[81] K. Deutsch, "Importance of information dominance," *Military Information Technology*, 2010.

[82] L. Stovall, "People, processes, and technology," *Military Information Technology*, 2010.

[83] L. Cumiford, "Situational awareness for cyber defense," in *2006 CCRTS: The State of the Art and the State of the Practice*, 2006.

[84] S. Jajodia and S. Noel, "Topological vulnerability analysis," in *Army Research Office Cyber Situational Awareness Workshop*, 2009.

[85] P. Cuviello and B. Kobel, "Cyber-awareness is a team sport," *Military Information Technology*, 2010.

[86] K. Condello, "Working together for real-time awareness," *Military Information Technology*, 2010.

[87] R. Koch and M. Golling, "Architecture for evaluating and correlating nids in real-world networks," in *5th International Conference on Cyber Conflict*, Tallinn, Estonia, 2013.

[88] O. McCusker, S. Brunza, and D. Dasgupta, "Deriving behavior primitives from aggregate network features using support vector machines," in *5th International Conference on Cyber Conflict*, Tallinn, Estonia, 2013.

[89] G. Conti, J. Nelson, and D. Raymond, "Towards a cyber common operating picture," in *5th International Conference on Cyber Conflict*, 2013.

[90] IBM. (2014, Feb) Analyst's notebook. IBM. [Online]. Available: http://www-03.ibm.com/software/products/en/analysts-notebook-family/

[91] Palantir. (2014, Feb) Palantir. [Online]. Available: https://www.palantir.com

[92] HP. (2014, Feb) Security information and event management. [Online]. Available: http://www8.hp.com/us/en/software-solutions/siem-arcsight/

[93] R. Xi, S. Jin, and X. Yun, "Cnssa: A comprehensive network security situational awareness system," in *IEEE 10th International Conference on Trust, Security, and Privacy in Computing and Communications*, Changsha, China, 2011.

[94] X. Yin, W. Yurcik, L. Yifan, K. Lakkaraju, and C. Abad, "Visflowconnect: Providing security situational awareness by visualizing network traffic flows," in *23rd IEEE International Conference on Performance, Computing, and Communications*, Phoeniz, AZ, 2004.

[95] S. Batsell, N. Rao, and M. Shankar. (2005) Distributed intrusion detection and attack containment for organizational cyber security. [Online]. Available: http://www.ioc.oml.gov

[96] W. Heinke, "What commander's need to know," *Military Information Technology*, 2010.

[97] M. Gregoire and L. Beaudoin, "The science of mission assurance," *Visualization and the Common Operational Picture*, 2005.

[98] I. Kotenko and A. Chechulin, "A cyber attack modeling and impact assessment framework," in *5th International Conference on Cyber Conflict*, Tallinn, Estonia, 2013.

[99] D. F. Vazquez, O. P. Acosta, S. Brown, E. Reid, and C. Spirito, "Conceptual framework for cyber defense information sharing within trust relationships," in *4th International Conference on Cyber Conflict*, Tallinn, Estonia, 2012.

[100] B. Casey. (2011, Mar) The IBM institute for advanced security blog. IBM. [Online]. Available: http://www.instituteforadvancedsecurity.com

[101] T. Pingel, "Key defensive terrain in cyberspace: A geographic perspective," in *Proceedings of the International Conference on Politics and Information Systems (PISTA)*, Orlando, FL, 2003.

[102] M. Sudit and A. Stoltz, "Information fusion engine for real-time decision-making (inferd): A perpetual system for cyber attack tracking," in *10th International Conference on Information Fusion*, Quebec, Canada, 2007.

[103] S. Yang, S. Byers, and J. Holsopple. (2008) Intrusion activity projection for cyber situational awareness. [Online]. Available: http;//www.ieeexplore.ieee.org

[104] M. Woodruff. (2014, Apr) The secret way companies are using big data to score you. [Online]. Available: http://finance.yahoo.com/news/the-secret-way-companies-are-using-big-data-to-score-you-135018683.html

[105] J. Kiss, "Inforgraphic: Where does Google get its revenue from?" *The Guardian*, 2011. [Online]. Available: http://www.theguardian.com/technology/pda/2011/jul/21/google-keyword-advertising

[106] J. C. Spang, "Open source information, social networking, and IC employees: An analytic vulnerability assessment," Master's thesis, National Defense Intelligence College, 2011.

[107] M. Drapeau and L. Wells, *Social Software and National Security: An Initial Net Assessment*, 2009.

[108] Department of Defense, *Directive-Type Memorandum 09-026 Responsible and Effective Use of Internet-Based Capabilities*, Feb 2010. [Online]. Available: http://www.defense.gov/NEWS/DTM2009-026.pdf

[109] ——, *DoD Operations Security (OPSEC) Program Instruction 5205-02E*, 2008. [Online]. Available: http://www.dtic.mil/whs/directives/corres/pdf/520502e.pdf

[110] ——, *DoD Operations Security (OPSEC) Program Manual 5205.02-M*, 2008. [Online]. Available: http://www.dtic.mil/whs/directives/corres/pdf/520502m.pdf

[111] Chief Information Officers Council. (2014, Oct) About. [Online]. Available: https://cio.gov/about/

[112] ——, *Guidelines for Secure Use of Social Media by Federal Departments and Agencies*, 2009. [Online]. Available: https://cio.gov/wp-content/uploads/downloads/2012/09/Guidelines-for-Secure-Use-Social-Media-v01-0.pdf

[113] ——, *Privacy Best Practices in Social Media*, 2013. [Online]. Available: https://cio.gov/wp-content/uploads/downloads/2013/07/Privacy-Best-Practices-for-Social-Media.pdf

[114] Federal Bureau of Investigation. (2009) Spear phishers: Angling to steal your financial info. [Online]. Available: http://www.fbi.gov/page2/april09/spearphishing_040109.html

[115] A. Felt and D. Evans, "Privacy protection for social networking APIs," in *W2SP*, Oakland, CA, 2008.

[116] Department of Defense, *DoDI 8550.01 DoD Internet Services and Internet-Based Capabilities*, 2012. [Online]. Available: http://www.dtic.mil/whs/directives/corres/pdf/855001p.pdf

[117] Office of the Chief of Public Affairs, *The United States Army Social Media Handbook Version 3.1*, 2013. [Online]. Available: http://www.slideshare.net/USArmySocialMedia/army-social-media-handbook-2012

[118] Division of Public Affairs, *The Social Corp, The U.S.M.C. Social Media Principles*, 2011. [Online]. Available: http://www.jbsa.af.mil/shared/media/document/AFD-120412-038.pdf

[119] Air Force Public Affairs Agency, "Air Force social media guide," 2013. [Online]. Available: http://www.af.mil/Portals/1/documents/SocialMediaGuide2013.pdf

[120] R. Mabus, *Internet-Based Capabilities Guidance: Unofficial Internet Posts*, 2010. [Online]. Available: http://www.public.navy.mil/bupers-npc/reference/messages/Documents/ALNAVS/ALN2010/ALN10057.txt

[121] M. Piesing, "Tweeting the Taliban: Social media's role in 21st century propaganda," *Wired*, 2012. [Online]. Available: http://www.wired.co.uk/news/archive/2012-03/20/military-social-media

[122] N. Ungerleider, "Inside the Israeli military's social media squad," *Fast Company*, 2012. [Online]. Available: http://www.fastcompany.com/3003305/inside-israeli-militarys-social-media-squad

[123] N. Perlroth, "2nd China Army unit implicated in online spying," *The New York Times*, 2014. [Online]. Available: http://www.nytimes.com/2014/06/10/technology/private-report-further-details-chinese-cyberattacks.html?_r=0

[124] Facebook Inc. (2014, Jul) Facebook. [Online]. Available: http://www.facebook.com

[125] Alexa. (2014, Jul) The top 500 sites on the web. [Online]. Available: http://www.alex.com/topsites

[126] Facebook Inc. (2014, Jul) Privacy. [Online]. Available: http://www.facebook.com/help/445588775451827

[127] D. Goodin, "Facebook page very much public, even when set as private," *The Register*, 2010. [Online]. Available: http://www.theregister.co.uk/2010/10/25/facebook_privacy_bypass/

[128] LinkedIn. (2014, Jul) What is linkedin? [Online]. Available: http://www.linkedin.com/static?key=what_is_linkedin

[129] ——. (2014, Jul) Compare account types. [Online]. Available: http://help.linkedin.com/app/answers/detail/a_id/71/ft/eng

[130] Office of Secretary of Defense. (2014, Mar) RDT&E budget item justification for defense operations security initiative. [Online]. Available: http://www.stratvocate.com/files/osd-p707/osd.html

[131] L. Sweeney, *Simple Demographics Often Identify People Uniquely*, 2000.

[132] Worldometers. (2014, Oct) World population. [Online]. Available: http://www.worldometers.info/world-population

[133] MathWorks. (2014, Aug) Statistics toolbox. [Online]. Available: http://www.mathworks.com/products/statistics/features.html#machine-learning

[134] Wikipedia. (2014, Sept) Naive bayes classifier. [Online]. Available: http://en.wikipedia.org/wiki/Naive_Bayes_classifier

[135] ——. (2014, Sept) K-nearest neighbors algorithm. [Online]. Available: http://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm

[136] ——. (2014, Sept) Homophily. [Online]. Available: http://en.wikipedia.org/wiki/homophily

[137] ——. (2014, Sept) Support vector machine. [Online]. Available: http://en.wikipedia.org/wiki/support_vector_machine

[138] ——. (2014, Sept) Random forest. [Online]. Available: http://en.wikipedia.org/wiki/random_forest

[139] L. K. Johnson, *Strategic Intelligence*. Greenwood Publishing Group, 2007.

[140] Kaspersky Lab, *The Evolution of Phishing Attack: 2011-2013*, 2013. [Online]. Available: http://media.kaspersky.com/pdf/Kaspersky_Lab_KSN_report_The-Evolution-of-Phishing-Attacks-2011-2013.pdf

[141] Department of the Army, *Army Web Risk Assessment Cell*, 2014. [Online]. Available: http://www.eff.org/files/filenode/070216EGS/031607_army_blog01.pdf

[142] Androit C4ISR Center, *The Impact of Internet Social Networking on the Vulnerability of United States Air Force Personnel to an Adversary Influence Operations*, 2007.

[143] L. Beckett. (2014, Jun) Everything we know about what data brokers know about you. [Online]. Available: http://www.propublica.org/article/everything-we-know-about-what-data-brokers-know-about-you

[144] Datalogix. (2014, Jul) Not all audiences are created equal. [Online]. Available: http://www.datalogix.com/audiences/

[145] Federal Trade Commission, *FTC Recommends Congress Require the Data Broker Industry to be More Transparent and Give Consumers Greater Control Over Their Personal Information*, 2014.

[146] B. McGarry. (2013, Mar) Army set to introduce smartphones into combat. [Online]. Available: http://www.military.com/daily-news/2013/03/27/army-set-to-introduce-smartphones-into-combat.html

[147] T. Suzuki. (2014, Feb) Ground troops may soon have smartphone-controlled drones. [Online]. Available: http://www.military1.com/army/article/460020-ground-troops-may-soon-have-smartphone-controlled-drones

[148] H. Leonard, "There will soon be one smartphone for every five people in the world," *Business Insider*, 2013. [Online]. Available: http://www.businessinsider.com/15-billion-smartphones-in-the-world-22013-2

[149] C. Osborne. (2013, Oct) Malicious apps, mobile malware reaches 1 million mark. [Online]. Available: http://www.zdnet.com/malicious-apps-mobile-malware-reaches-1-million-mark-7000021371/

[150] J. Gilbert. (2012, Feb) iPhone app privacy: Path, Facebook, Twitter, and Apple under scrutiny for address book. Huffington Post. [Online]. Available: http://www.huffingtonpost.com/2012/02/15/iphone-privacy-app-path-facebook-twitter-apple_n_1279497.html

[151] L. Constantin. (2014, Feb) New iOS flaw allows malicious apps to record touch screen presses. [Online]. Available: http://www.pcworld.com/article/2101580/new-ios-flaw-allows-malicious-apps-to-record-touch-screen-presses.html

[152] T. Book and D. S. Wallach, "Longitudinal analysis of android ad library permissions," in *MSoT*, San Francisco, CA, 2013.

[153] ——, "A case of collusion: A study of the interface between ad libraries and their apps," in *3rd Annual ACM CCS Workshop on Security and Privacy in Smartphones and Mobile Devices*, Berlin, Germany, 2013.

[154] J. Gantz and D. Reinsel, "The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east," *International Data Corporation*, 2012. [Online]. Available: http://www.emc.com/leadership/digital-universe/2012iview/index.htm

[155] J. Winter. (2014, Sept) Law enforcement bulletin warned of ISIS urging jihad aattack of us soil. Fox News. [Online]. Available: http://www.foxnews.com/world/2014/09/17/law-enforcement-bulletin-warned-isis-urging-jihad-attacks-on-us-soil/