Center for Cyber
& Homeland Security
THE GEORGE WASHINGTON UNIVERSITY

Trends in Technology and Digital Security

Methods of Analysis and the Utility of New Tools for Threat Forecasting

Issue Brief 1 in a Series
Based on Fall 2017 Symposium Proceedings

Speakers/Panelists
Charlie Allen - The Chertoff Group
Mike Davis - CounterTack
Sean Kanuck - International Institute for Strategic Studies
Jason Matheny - Intelligence Advanced Research Projects Agency
Teresa Shea - In-Q-Tel

Panel Moderator
Frank J. Cilluffo

Panel Rapporteur
Sharon L. Cardash

This publication is the exclusive work product of the Center for Cyber & Homeland Security. It was made possible thanks to the financial support of Razor's Edge Ventures and Raytheon Company.

RAZOR'S EDGE

Raytheon

Issue Brief Series on Trends in Technology and Digital Security
Methods of Analysis and the Utility of New Tools for Threat Forecasting

On September 14, 2017, CCHS convened a Symposium on Trends in Technology and Digital Security. Four panels addressed emerging threats and their implications for security policy, with a focus on digital infrastructure protection and anticipatory analysis. In a series of Issue Briefs, CCHS shares the findings and recommendations that emerged from the Symposium, primarily on a not-for-attribution basis. This first Brief in the series addresses Methods of Analysis and the Utility of New Tools for Threat Forecasting.

The Threat Climate: National Security at a Time of Rapid Technological Change

The current pace of technological change is striking. The world has more engineers than ever before and there are institutional mechanisms such as peer review that allow individual discoveries to be incorporated into technology much faster than ever before. The dark side to this is that there are more technologies to worry about now than ever before. Think advanced cyber weapons, drone swarms, and synthetic viruses—none of which could be anticipated 70 years ago. This accumulation of risk includes offensive asymmetric opportunities that render defensive systems disproportionately costly, disproportionately effective, or only temporary in their effectiveness.

Perhaps most significant as a driver of these trends is the share of commercial innovation relative to that of government, which means that the most disruptive technologies are now often publicly available in a way that was not true before the 1960s. In addition, the period of superiority for the U.S. government's defense science and technological innovation is shrinking over time. What used to be a period of 5 to 10 years of superiority or dominance in areas of technology in which the government invested is now 6 months to 1 year, if that—and there are even some areas where the government may lag behind industry (at least its median level).

This shift to a world in which most technology is publicly available and commercially funded makes for a technological environment that is increasingly more complicated and consequential. From a U.S. government standpoint, in general, the focus ought to be on the threats that pose existential risks to the country—such as nuclear war, electromagnetic pulse (EMP), cyberattacks that permanently cripple infrastructure, a large-scale pandemic, and certain emerging threats posed by biology.

Consider the especially worrisome aspects of biology. The difficulty of controlling even naturally occurring diseases like malaria or tuberculosis—even when massive resources are applied—is sobering. What if intelligent adversaries were applying some ingenuity to creating diseases? Those organisms can be weaponized effectively to spread efficiently throughout populations—evolving not by random mutations, but by engineering principles. While most of this type of knowledge used to be locked up in large-scale national programs; that changed in 2003, when the first virus was synthesized from scratch. No longer did you need a sample of the organism; instead all you needed was the sequence of DNA or RNA that

would encode the virus, and the raw chemicals to turn that code into biology. Biology was thus transformed into computer science or an engineering discipline.

In fact just last month, a Canadian scientist using commercially available equipment and chemicals, synthesized from scratch at a cost of $100,000 the first pox virus. This kind of technology is getting cheaper at double the rate of Moore's law, and it will get more sophisticated with the introduction of tools like CRISPR (which is used for gene editing) or with the development of things like gene drives which allow particular genes to become prevalent in a population very quickly. The upshot is that a misanthrope with biology training could, for example, re-create smallpox in his or her basement. Another potent illustration: the blueprint for the influenza virus is publicly posted on an NIH website and could now be re-created for $1,000,000. That virus is more effective than a hydrogen bomb in terms of mortality. (A naturally occurring influenza outbreak killed 100 million people worldwide in 12 months, about a century ago; it was the most statistically significant mortality event in human history).

As the requisite skills and budgets continue to shrink in accessibility, biological threats present a particularly challenging national intelligence problem. It is harder to detect distinct signatures of a biological weapons effort, and the set of actors that we have to worry about is large. Individuals are especially challenging as they leave a smaller digital footprint (than nation-states or groups) and have a broader set of motivations—including motivations are not subject to deterrence. Apart from the malicious actor, with powerful technologies, accidents can become especially catastrophic and kill tens of millions of people, even if they are infrequent.

Against this background, we need to think about new phenomenology that can help to reveal low-signature, dual use activities. We need improved measurement and signature intelligence (MASINT) for being able to assess chemical, biological, radiological, and nuclear (CBRN) activities from standoff distances. We need to broadly strengthen U.S. capabilities for scientific and technical intelligence. This list is merely illustrative and not comprehensive. Many of the tools that we have focus on looking at publications or patent trends, but we also need tools that look at what is increasingly becoming the playing field for emerging science which is conferences or social media. We also need to find new tools that are able to bring down the analytic burden for our relatively small number of S&T analysts.

The initiative will always remain with the attacker, but we should prepare for surprises even if cannot prevent them all.

Forecasting Threat: Methods and Tools

The mission of the intelligence community (IC) is to avoid surprise—to understand threats, see them early, and take action. Yet the historical record is checkered. The Arab Spring, the so-called caliphate declared by the Islamic State, and Russia's move into the Crimea all came as a surprise to, rather than a warning from, the IC. Looking ahead however, there is a transformation occurring in the form of digital information and big data analytics that we

can bring to bear in the field of anticipatory intelligence. We now have great opportunity to gather, store, and analyze more information than ever before. In terms of tools and equipment, consider the Oak Ridge supercomputer that will be up and fully running in less than five years; it will be the most powerful in the world, second to none.

The National Intelligence Strategy put out in 2014 under then-Director of National Intelligence Clapper discussed anticipatory intelligence and the need for having a cross-cutting mission to make sure that we get "left of boom"/left of surprise. As one panelist stated, "My view is that we're a long way from that." While the country trains people in such a way that when we go to war we do it better than anyone in the world, there is no analogue for anticipatory analysis or warning—it is not taught at the CIA's Kent School, nor is it taught in requisite detail at the National Intelligence University.

One area that we have clearly not done well at is threat forecasting in cybersecurity. In this context, you can think about threat forecasting in an architectural framework that is composed of three pieces: digital intelligence (collection and analysis); trusted infrastructure (countermeasures and the analysis you apply); and an underpinning by analytic workflow and the ability to do agile operations. Technologies will drive threat intelligence and threat forecasting. Machine learning, a subset of artificial intelligence, is being applied today in: behavioral analytics, especially the insider threat; situational awareness of your networks, to identify things that are happening on your networks in real-time, thereby giving you things to take action against; and threat intelligence, automating rote tasks that the IC's analytical staff spend a lot of time on.

Threat intelligence has to be both actionable and timely. There is a difference between threat data and information, and threat intelligence. Analysts are overwhelmed with the amount of data and information that they are getting. They are consumed by trying to correlate these disparate pieces of information to provide context around threats. A lot of this work can be done today with machine learning which results in threat intelligence. However this intelligence has to be tailored to your particular environment and this is where we fall down.

Consider threat forecasting in the cyber context. Tailoring the intelligence to that environment means that you have to be able to answer the questions: What are my digital assets? And which ones are most important? We tend to treat all the data the same, but not all the data is the same. Really knowing what are those priority assets (or priority information or priority intellectual property) that we actually want to protect, allows us to prioritize vulnerabilities and have cybersecurity analysts focus on those top-tier threats against your highest risks. In turn, that enables faster action against threats, faster decision-making on what is being seen, and discovery of new threats you might not have known were out there.

It is important to think through the goal of the action you want to take, based on threat intelligence. There are at least three levels at play here. The first is strategic. In this instance, you care about attribution, e.g., is this an advanced persistent threat (APT)? If so, they will not give up and have lots of resources. Think about what you will do with that information. Second is the operational level. This is where tactics, techniques, and procedures (TTPs) are

coming at you. These must be understood in order to address your vulnerabilities—not just in the network and in the equipment—but also those individuals susceptible to clicking on emails, files, or phishing schemes (that keep getting better and better); people are always the greatest weakness factor. Third is the tactical level. Here the relevant questions are: Do I really have all my vulnerabilities patched? What are those vulnerabilities? And where do I apply my precious cyber resources?

As one panelist put it, "You don't have a needle in a haystack problem. You have a needle in a needle-stack problem. There are all kinds of threats (biological, nanotechnology, etc.); so it's the analysis problem that's really plaguing us right now." And that problem, in turn, breaks into two parts. First, we do have a lot of data, yet analysts are still not able to get the information they actually need to make a decision in a timely manner, especially from a cyber perspective. Second, the analysts themselves are not even looking in the right spots: for many years we have been ignoring the workstation, laptops, and servers in our environment, while almost all of our effort has been on the network. This is the equivalent of going to a crime scene investigation for a murder in a home and not being able to leave the street. It is only in the past two years that we have been looking at the endpoint—the actual systems we use, that connect to the assets that are actually critical.

There is also a cross-domain issue with cybersecurity. What happens if we have a Stuxnet from the biological perspective, meaning that somebody thinks they are making a legitimate virus that could actually help us, and unknowingly the software is actually creating a bad virus underneath them? When you blend these two domains (biology and information security) together you start getting a situation for which analysts are not ready. And when it comes to training at information security in cyberspace, we are good at training in one domain (such as network, or threat intelligence); but we are not good at bridging domains, or looking at information from each of several domains and then applying it.

With threat forecasting and threat analysis, there are certainly some great opportunities around machine learning and big data; but these opportunities are double-edged: already attackers are using machine learning to improve their phishing campaigns. The same things that work to better target advertising are being used to better target phishing. The adversary is getting just as smart in terms of the cost and speed of their innovation; and this far outpaces our defenses because we just do not have the right type of training or technology in place to look at the right type of data.

What is the best use of the assets in the intelligence community? Bear in mind, for instance, that some of our private sector companies are now able to do forensic attribution of cyber incidents as well or better than IC assets (except where they have very particular secret accesses). The template of knowns and unknowns is useful to invoke here in answer to the question. First consider the known unknowns. These are the frameable questions such as who will win the next election in Country X? Humans are not gifted at probability analysis or at predicting these things, yet we are finding increasingly that the technology we are creating is much better at that. Consider big data analysis of sentiments on social media, or meme propagation where you want to infect or influence behavioral patterns. Therefore this may not be the best place for our human analysts, since we are automating it better and better.

Second, consider the unknown knowns. This is where you are looking for specific indicators such as which would-be terrorists are talking about a bomb event. This is where you know what you are looking for, but you have just inordinately large data sets that you need to find the datum in. That data culling again needs to be automated, and that is where we are progressing towards. To put this in a cyber context, this is what red teams and hunt teams need to be doing. The red teams are looking for vulnerabilities, and the hunt teams are looking for indirect evidence of a previous compromise. Again, this needs to get increasingly automated.

Third, consider the unknown unknowns. Here the question is what do I need to be worried about? More specifically, what is happening anomalously in the contextual environment I am in, for which I am unprepared? This is the domain of genuinely anticipatory intelligence. And it is complicated by the fact that in previous decades and centuries, we were dealing with disruptive applications of known technology such as radar; whereas today it is the technologies themselves that are disruptive. This is a huge difference, brought on by the time horizon cycle of technological innovation. For example, if you were supposed to be doing strategic forecasting on the Arab Spring five years out, you would not even know the term "social media" let alone be in a position to assess the impact of social media on it—because Facebook had just left Harvard's campus a year earlier, and Twitter wouldn't be invented until 2006.

Where does this leave us when trying to consider existential risks in strategic forecasting? You have to turn the lens, instead of on the threat actors, you need to turn it introspectively—look at your society, your assets, and what are your critical digital assets that need protecting. Think about the Saudi Aramco cyber event a few years back. It devastated their corporate networks, but it did not get to and harm their production and transmission networks, the core assets of the enterprise. One could debate whether that was luck or design or prevention; but the important point is that Saudi Aramco suffered a huge price-tag but it was one they could weather and survive.

Coming back to biology, the most critical database that we need to be worried about is the human genome, and external efforts to undermine the integrity of that database. And it is not just information security coupled with biology; you also have to add in nanotechnology, because just as we are becoming increasingly able to undermine the integrity of organic platforms through molecular-level production and material science, you can also undermine the integrity of (or introduce unwanted additional features into) inorganic materials and platforms on which we rely in our daily life. Returning to the question of where do we need to spend those very limited critical human assets in the intelligence community, on things we cannot yet automate; that is the strategic forecasting question—and to get the most bang for your buck, you need to look introspectively at your society and your crown jewel assets that are based in carbon, because that is ultimately where your risks lie.

Specially-engineered crops that cannot be digested; or person-specific pathogens designed to go after particular races, or political leaders. This is no longer the stuff of fiction or wild imagination. It is in our best interests to seek to shape the future. There is plenty of scientific and technological talent worldwide, not all of it in friendly settings. For instance, how many nuclear scientists and cryptologists from the Former Soviet Union ended up in

North Korea and Iran? Meeting the national security challenges of the day requires better training and more exquisite tradecraft than we have today within the intelligence community. And until information sharing is at a higher quality, it will be hard to get better decisions. But the lessons of the past are clear: we have to be able to share information and break down bureaucratic silos—because historically, when this has not happened, we have been surprised.

Although our thinking tends to concentrate upon adversaries and malicious actors, the technologies that exist today and that continue to develop rapidly are extraordinarily powerful, and the possibility for technological accidents that are catastrophic is vast—so much so that one participant estimated that humanity has only a 70% probability of making it through the next century.

About Us

The Center for Cyber & Homeland Security (CCHS) at the George Washington University is a nonpartisan "think and do" tank whose mission is to carry out policy-relevant research and analysis on homeland security, counterterrorism, and cybersecurity issues.

Website http://cchs.gwu.edu **Email** cchs@email.gwu.edu **Twitter** @gwcchs