ARL-SR-0443 ● Apr 2021

Final Report and Recommendations of the North Atlantic Treaty Organization (NATO) Research Task Group IST-129 on Predictive Analysis of Adversarial Cyber Behavior

by Dennis McCallam, Tracy Braun, Bernt Akesson, David Aspinall, Roman Faganel, Heiko Guenther, Matthew Kellet, Joseph LoPiccolo, Peeter Lorents, Wim Mees, Juha-Pekka Nikkarila, Teodor Sommestad, and Margaret Varga

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

ARL-SR-0443 ● APR 2021

Final Report and Recommendations of the North Atlantic Treaty Organization (NATO) Research Task Group IST-129 on Predictive Analysis of Adversarial Cyber Behavior

Dennis McCallam
US Naval Academy and George Mason University

Tracy Braun
Office of the Director, DEVCOM Army Research Laboratory

Bernt Akesson and Juha-Pekka Nikkarila, *Finnish Defence Research Agency*

David Aspinall, *University of Edinburgh*

Roman Faganel, *Slovenia Ministry of Defence*

Heiko Guenther, *Fraunhofer Institute for Communication, Information Processing and Ergonomics (FKIE)*

Matthew Kellet, *Defence Research and Development Canada*

Joseph LoPiccolo, *US Naval Postgraduate School*

Peeter Lorents, *Estonian Business School*

Wim Mees, *Royal Military Academy*

Teodor Sommestad, *Swedish Defence Research Agency*

Margaret Varga, *Seetru Ltd and Oxford University*

Approved for public release; distribution unlimited.

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*
April 2021

2. REPORT TYPE
Special Report

3. DATES COVERED (From - To)
1 October 2015–30 April 2019

4. TITLE AND SUBTITLE
Final Report and Recommendations of the North Atlantic Treaty Organization (NATO) Research Task Group IST-129 on Predictive Analysis of Adversarial Cyber Behavior

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Dennis McCallam, Tracy Braun, Bernt Akesson, David Aspinall, Roman Faganel, Heiko Guenther, Matthew Kellet, Joseph LoPiccolo, Peeter Lorents, Wim Mees, Juha-Pekka Nikkarila, Teodor Sommestad, and Margaret Varga

5d. PROJECT NUMBER
NATO IST-129

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
DEVCOM Army Research Laboratory
ATTN: FCDD-RLD-FT
Adelphi, MD 20783-1138

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-SR-0443

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
NATO Science and Technology Organisation
Collaboration Support Office (CSO)
BP 25, 92201 Neuilly sur Seine, France

10. SPONSOR/MONITOR'S ACRONYM(S)
NATO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This report summarizes the work and findings of the North Atlantic Treaty Organization (NATO) Research Task Group (RTG), Information Systems Technology (IST)-129, on Predictive Analysis of Adversarial Cyber Operations. The RTG found overall there was little in the way of direct research and solutions of predicting a cyber-adversary who launches an attack against a known vulnerability with an unknown exploit. As such, the work of IST-129 contains a body of work that provides researchers and organizations a point of departure for continuing research. Of all our many findings and recommendations, the most important is that prediction of adversarial operations in cyberspace is complex, but can be decomposed. Prediction offers great potential in many areas of cyber defense. Predicting adversarial operations will be a multimethod approach. A common taxonomy both for and about the threat, along with machine-readable language, will help. Cyber defense itself needs to be protected. Modelling of closed network systems is needed and we need data sets that are representative of reality.

15. SUBJECT TERMS
predictive analysis, predictive analytics, ensemble modeling, adversarial cyber operations, cyber situational awareness, cyber defense

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Tracy Braun
a. REPORT	b. ABSTRACT	c. THIS PAGE	UU	95	19b. TELEPHONE NUMBER (Include area code)
Unclassified	Unclassified	Unclassified			301-394-4954

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Contents

List of Figures

List of Tables

Acknowledgments

The work of the Information Systems Technology (IST)-129 Research Task Group (RTG) was performed by researchers who had the support of their respective organizations. On behalf of the IST Panel, we would like to express gratitude to the following organizations for their support of the key researchers. The IST-129 RTG was composed of Dr Dennis McCallam (Chair), Adjunct Professor at US Naval Academy and George Mason University, United States; Lt Cdr (Eng.) Dr Bernt Akesson, Finnish Defence Research Agency; Prof David Aspinall, University of Edinburgh, United Kingdom; Dr Tracy Braun, US Army Combat Capabilities Development Command Army Research Laboratory, United States; Roman Faganel, MSc, Slovenia Ministry of Defence, Slovenia; Heiko Guenther, Fraunhofer Institute for Communication, Information Processing and Ergonomics (FKIE), Germany; Dr Matthew Kellet, Defence Research and Development Canada, Canada; Joseph LoPiccolo, US Naval Postgraduate School, United States; Prof Peeter Lorents, Estonian Business School, Estonia; Dr Wim Mees, Royal Military Academy, Belgium; Capt (Eng.); Dr Juha-Pekka Nikkarila, Finnish Defence Research Agency, Finland; Dr Teodor Sommestad, Swedish Defence Research Agency (FOI), Sweden; and Dr Margaret Varga, Seetru Ltd and Oxford University, United Kingdom.

1. Introduction and Key Findings

1.1 Introduction

This report summarizes the work and findings of the North Atlantic Treaty Organization (NATO) Research Task Group (RTG), Information Systems Technology (IST)-129, on Predictive Analysis of Adversarial Cyber Operations. The work of this RTG was initiated in late 2015 and completed in April 2019. This report is unclassified and open to NATO nations, Partner for Peace nations, Mediterranean Dialogue, Istanbul Cooperation Initiative nations, and Global Partners. The RTG Chair was Dr Dennis McCallam, formerly with Northrop Grumman Corporation, and currently with the US Naval Academy and George Mason University.

The RTG members were Lt Cdr (Eng.) Dr Bernt Akesson, Finnish Defence Research Agency; Prof David Aspinall, University of Edinburgh; Dr Tracy Braun, US Army Combat Capabilities Development Command Army Research Laboratory; Roman Faganel, MSc., Slovenia Ministry of Defence; Heiko Guenther, Fraunhofer Institute for Communication, Information Processing and Ergonomics (FKIE); Dr Matthew Kellet, Defence Research and Development Canada; Joseph LoPiccolo, US Naval Postgraduate School; Prof Peeter Lorents, Estonian Business School; Dr Wim Mees, Royal Military Academy; Capt (Eng.) Dr Juha-Pekka Nikkarila, Finnish Defence Research Agency; Dr Teodor Sommestad, Swedish Defence Research Agency (FOI); and Dr Margaret Varga, Seetru Ltd and Oxford University.

The work of the RTG represents one of the initial, if not *the* initial, attempts at organizing at an international level the evaluation of prior research into predicting cyber events. The RTG found overall there was little in the way of direct research and solutions of predicting a cyber-adversary who launches an attack against a known vulnerability with an unknown exploit. As such, the work of IST-129 contains a body of work that will provide researchers and organizations a point of departure for continuing research.

1.2 Important Results and Findings

We had many findings and recommendations that are presented in this final report. While we present our detailed findings from the NATO RTG's Specialist Meeting and the overall RTG in both Sections 10 and 12, respectively, here is a summary of the most important of those findings.

1) **Prediction of adversarial operations in cyberspace is complex, but can be decomposed**. Based on all the research and investigation, the RTG concluded that at this time predicting an adversary's next action with absolute certainty is difficult. However, we may be able to predict likelihood (plausibility) of the next adversarial operation from a fixed number of options. For example, threat modeling using attack graphs or similar methods can reduce the space of uncertainty in predicting possible attack routes. This is an important conclusion since it properly characterizes the current state of the practice.

2) **Prediction offers great potential in many areas of cyber defense.** Predicting the likelihood of potential next adversarial actions could be effectively used more widely in cyber defense. Key areas that could benefit include prioritization of patching and remedial activities, enhancing defense courses of action, allocation of analytical resources (human and machine), and reprioritization of mission resources.

3) **Predicting adversarial operations will be a multimethod approach.** Detecting known exploits has been solved (using classic statistical approaches), but prediction is still a challenge. Multiple analytical methods will be needed, not just to predict based on previous data but also to take into account future possibilities that have not yet been seen—extending threat prediction in the (known, unknown) region to the (unknown, unknown) space. Data sets relevant to this are necessarily smaller, incomplete, and extremely hard to come by.

4) **A common taxonomy both for and about the threat, along with machine-readable language, will help**. Recognizing the means of facilitating communication across domains and borders is essential to be able to share information on adversarial operations. We recommend the use of Structured Threat Information eXpression (STIX; Barnum 2012) and Trusted Automated Exchange of Indicator Information (TAXII; Connolly et al. 2014) as a means of standardizing adversarial information communication. In addition, we also recommend using capability descriptions of the threat to provide insight into an adversary's skill set. The OASIS open standard STIX lends itself to more efficient communications across all entities working the cyber-event prediction problem and will be particularly useful as an interchange format as tools emerge, for example, to combine multiple predictions into dashboards.

5) **Cyber defense itself needs to be protected**. The committee noted that prediction approaches and analytics constitute the cornerstone of defensive

cyber operations, and as such, need to be made resilient to cyber deception or manipulation, which is critical to maintain trusted operations. This area certainly needs further research.

6) **Modeling of closed network systems is needed**. The goal of prediction of adversarial operations within closed networks is more complex: new attack vectors and command and control structures for closed and controlled internet borders need to be modeled. In addition, this needs to be extended to recommendations for cyber defense against adversaries who have their base of operations behind closed networks.

7) **We need data sets that are representative of reality**. There is a lack of clear benchmarks for emerging technology; this implies an inability to compare solutions on how things should perform due to a lack of agreed measurements and assessments that represent ground truth.

2. Review of Technical Activity Proposal (TAP) Compliance

The work of the IST-129 RTG was governed by the original Technical Activity Proposal (TAP) (NATO 2015), which had four primary tasks. This section provides a detailed discussion of the IST-129 RTG's work, but in essence, all four objectives of the original TAP were successfully met. More importantly, with the publishing of this report, we will have successfully published several papers on this topic and have others in prepublication, which adds greatly to a thin body of knowledge (McCallam et al. 2018, 2019, 2021). Throughout the research, the RTG continually evaluated the TAP to ensure our output would meet the desired outcomes. This final report documents the work we performed and will be available to support NATO and its members in future research and planning activities. Each of the goals within the TAP are discussed separately. The four overall goals were the following:

1) "To characterise the current state of research in the field of Predictive Analysis of Adversarial Cyber Operations and develop a prioritised assessment of potential methodological and technical approaches with the focus on intelligence preparation of the cyber battlefield." This goal is split into two discussions: Section 2.1 characterizes the current research and Section 2.2 discusses the intelligence preparation of the battlefield.

2) The next goal, to "articulate the similarities and differences with conventional warfare approaches to the current Predictive Analysis of Adversarial Courses of Action (COAs)," is discussed in Section 2.3.

3) The goal of a "focused technical workshop at the NATO Unclassified (NU) level" was designed to assess and validate the current state of the art in the

academic, defense, and other communities. The Specialist Meeting is summarized in Section 2.4 and discussed in detail in Section 10.

4) Finally, the RTG was "to develop an initial roadmap for development of a comprehensive set of methodologies, technologies and tools for advancing the pro-active Predictive Analysis of Adversarial Cyber Operations." This is discussed in Section 2.5.

2.1 Characterize Current Research

Characterizing the current state of research was straightforward. Basically, little to nothing has been done on the strict problem of prediction. And by prediction, the committee took the definition to mean identification of the next event as opposed to which of the next events is likely. That is a critical distinction. Within the first goal, we have several results and observations.

With respect to the current state of research in cyber-operations prediction, we found little to nothing in terms of pure predictive analysis. And by that, we mean prediction of the next event and not picking the most-likely event from a set of possible events. That would be more in line with knowing perhaps some or all of the previous events by this adversary and then postulating a set of next steps based on the previous event set matching a set of known and defined adversary techniques, tactics, and processes (TTPs).

We did find data in *identification* of cyber events, but not in the area of linking current events to potential future events. In Section 11.5, we present various ways we examined the body of knowledge to identify relevant research on predictive analysis in cyberspace. But our observation was that little work has been accomplished on the exact problem of predicting the next move of an adversary in cyberspace. This underscores the importance of the committee's work in that this represents a starting point for future research and a comprehensive analysis of the current state of the art.

Developing the prioritized assessment of potential methodologies led us to recommend a series of follow-on activities, since our research showed we are still in early stages of understanding and characterizing the problem. What we did not recommend directly was a follow-on to IST-129 in its current form. The series of follow-on activities we recommend are as follows:

- **RTG designing an experiment using combinatorial analytics for predictability**. One of the observations made by IST-129 was that any analytic or process that could be used to predict adversarial behavior will not be a single algorithm. It will more likely be a compound analytic that is more Observe, Orient, Decide, Act (OODA) loop organized (Brumley et al. 2006; Sorensen 2010). This activity would evaluate candidate compound architectures.

- **Exploratory team (ET) evaluating the viability of using TTPs as a means of predicting adversarial behavior**. Another observation made by IST-129 was that adversarial behavior could be evaluated by looking at known TTPs and then extrapolating after so many events to predict the next event, making the assumption that events are linearly dependent. This thought is not without caution since another observation made is that next events could also be thought of as linearly independent. The concept has some level of usefulness and IST-129 believes an ET could provide some scientific support for or against such an approach.

- **An RTG created to predict adversarial behavior that is based on attack surface analysis expansion**. Another observation from IST-129 was that predicting behavior for threats using known vulnerabilities with unknown exploits could be accomplished by expanding attack trees, which reduces the problem set from one of prediction to one of pattern matching. The intent of this RTG would be to develop those approaches to expand attack trees around known vulnerabilities to encompass unknown exploits.

- **Experimentation into identifying adversarial behavior based on TTP and attack tree expansion.** This would involve designing and performing an experiment that would combine the previous two ideas. It is expected that this would be a limited experiment and one that could almost be a tabletop red team approach (Yuen et al. 2015), but that the results of such an experiment could provide deeper insight into *how* to use predictability approaches.

The current research investigation highlighted an area that needs to be addressed in a number of areas beyond predictability. This is the data set deficiency for evaluation of any potential solution. The lack of a data set that represents real traffic in either commercial or military environments not only hampers solution testing, but prevents benchmarking proposed solution effectiveness.

While not a complete or recommended list, the committee at various times referred to some of the commercially available products to help identify what data could be available, the format of the data, and any processing that currently exists: Norse, Checkpoint SW, FireEye, Arbor Networks, Trend Micro, Akamai, Fortinet, Splunk, and ArcSight.

2.2 Intelligence Preparation of the Battlefield

Communication within military organizations is challenging in and of itself, but most operations have defined specific ways and messages to ensure minimal or no miscommunication. For cyber, this is not necessarily the case at this point in time. Further compounding the problem is the lack of extensive taxonomy for cyber operations. This impedes any intelligence preparation of the battlefield.

During research, the RTG found several areas where communications could be improved and recommends that strong and consistent communication approaches be used to share information. Specifically, these are in areas conveying event information and describing the capability levels of the threat. To this end, we cite STIX (Barnum 2012) and TAXII (Connolly et al. 2014) as a means of characterizing and communicating event information, and the Defense Science Board (DSB) report (Gosler and Von Thaer 2013) in Table 1 and the US Government Accountability Office (GAO) report (GAO 2018) in Table 2 to communicate threat capability level.

The six-level, three-tier model described in the DSB report (Table 1) is directed at describing the cyber-adversary with respect to their depth of knowledge and capabilities. It is basic in its approach to categorizing the threat and more appropriately sets by level the variety of tactics an adversary would be able to employ. The four-tier model in Table 2 provides deeper insight into the specific approaches an adversary would employ. In either case, the observation from the RTG is that indicating "this is a DSB Tier III threat" provides deep context into the skill level of an adversary.

Table 1 Description of cyber threats with respect to vulnerability vs. exploit

Threat tier	Adversary capability	Vulnerability	Exploit	Response
I	Uses known exploits against known vulnerabilities	Known	Known	Known: usually a patch
II	Develops tools from and for publicly known vulnerabilities			
III	Discover/use unknown malicious code to steal/modify data	Mostly known, an existing vulnerability being exploited	Variant of known or new exploit not recognized by signature analysis; requires other forms of analysis	Not developed or known
IV	Discover new vulnerabilities and develop exploits			
V	Create exploitable vulnerabilities in products for networks/systems	Unknown and unidentified. Previously not known or identified.	Unknown and previously unseen (since function of unseen/unknown vulnerability)	Completely unknown and may impact system architecture
VI	Execute full-spectrum cyber operations and apply at scale			

Note: Derived from p. 22 of the DSB report (Gosler and Von Thaer 2013).

Table 2 Updated definition of cyber threats based on capability

Threat type	Description
Nascent	Little-to-no organized cyber capabilities, with no knowledge of a network's underlying systems or industry beyond publicly available open-source information.
Limited	Able to identify—and target for espionage or attack—easily accessible unencrypted networks running common operating systems using publicly available tools. Possesses some limited strategic planning.
Moderate	Able to use customized malware to conduct wide-ranging intelligence collection operations, gain access to more isolated networks, and in some cases create limited effects against defense critical infrastructure networks.
Advanced	May conduct complex, long-term cyber-attack operations that combine multiple intelligence sources to obtain access to high-value networks. May develop detailed technical and system knowledge of the target system to deploy more damaging cyber-attacks.

Note: Taken from GAO-19-128, Weapon systems cybersecurity, p. 9 (GAO 2018).

There has been a communication issue in terms of how different observers characterize cyber events and cyber-adversaries. Our recommendation is that the communication issue can be solved by using a consistent means of characterizing and sharing information. We believe that the STIX concept of describing cyber events provides the desired consistency and conciseness for proper intelligence on observed events. Furthermore, characterization as to the capabilities of the adversary defined by either the DSB or through the GAO report (GAO 2018) become important and prevent underestimating the space of potential "next events". This also supports the desire to communicate consistently and concisely. The committee recommends that cyber-threat discussions leverage and use STIX and the two threat capability descriptions (in Tables 1 and 2) to be consistent and clear.

2.3 Similarities and Differences with Conventional Warfare

Our next goal was to articulate the similarities and differences with conventional warfare approaches to the current predictive analysis of adversarial COAs. Our research uncovered some similarities along with a key difference. Being able to predict next steps relies on knowing where you currently are, and this applies to both conventional and cyber warfare. In addition, this situational awareness extends to knowing where the lines of defense are, what areas are being protected, and what areas are not.

We identified some similarities and differences, but the key observation is in the temporal domain. Conventional warfare has temporal and physical restrictions in movement of assets, while in the cyber domain there is no restriction of movement. In addition, conventional warfare has known observables that can aid in determining intent, whereas cyber warfare has not matured to the level of understanding or identifying precursor observables. The discussions about those differences are summarized in Table 3.

Table 3 Results of cyber and physical warfare similarities discussion

Comparison categories	Physical or conventional warfare	Cyber warfare
How is the target identified?	Usually a high-value strategic asset (from individual to system to geography). In most cases, the target is chosen as a proportional response for adversary action. Can be preemptive.	Many reasons why target selected (hacktivism, disruption, theft, etc.). Usually research and identification through open source, system probing, and collateral analysis.
What are the choices and types of weapons?	Wide choice of weapons and effects. Can be conventional, nuclear, chemical, biological, jamming, etc. Dependent upon the desired effect.	Many forms of malware exist: worm, virus, malware, Trojan, specialized (Stuxnet), etc.
What are the forms of delivery for the chosen weapon?	Delivery mechanisms are typically multiuse but constructed with respect to range (missile, shell, bullet, jammer, laser, etc.).	Wide variety of available delivery means: Email, website, USB, patch, CD, game/music, related software, etc.
What are the guidance characteristics?	Typically geospatial and requires a physical "address" where the target of interest resides. Guidance can also be done via secondary means, laser for example.	Specifically directed to exploit vulnerabilities that could be caused by adversary. Looking for privileged or specific pieces of software or system functionality.
What are the issues in attribution of the attack?	Attribution: where easier than who. Not so anonymous.	Difficult and highly complex. Closed networks provide additional cover for adversary.
What is the physical effect of the attack?	Surgical or related destruction of target from precision standpoint.	Most times no direct physical symptom, although cases (like Stuxnet) can provide derivative physical damage. Usually involves "ownership" of the target resulting in loss of trust with the target system. Also possibility to exfiltrate/modify information.

Conventional warfare approaches rely heavily on the laws of physics and known information about current locations of both friendly and adversarial forces. In terms of similarities, the knowledge of where an adversary exists in both physical space and cyberspace are important. For example, if an adversarial aircraft is moving toward a target, the prediction of how long that will take relies heavily on the laws of physics and aerodynamics along with range rate and direction. Specifically, a plane at time T flying at Mach 1 cannot instantaneously show up at time T+1 in a location that is on the other side of the world. Cyber events do not have current position and speed as a restriction. If one considers cyberspace in Cartesian coordinate space, it is possible for two events to occur either sequentially or simultaneously at opposite sides of that coordinate space. That same coordinate space applied to conventional warfare would have the events occurring only where the laws of dynamics are obeyed. One way to predict what happens next in cyberspace and what is its credibility level is to rely on numerical evaluation of similarity for situations and developments. Our observation is that current conventional warfare has well-defined observables that indicate battlefield preparation. Cyberspace does not have established precursors and routinely we cannot identify a cyber event until or after it occurs. There is far more predictability in where a conventional attack will occur and in many cases, when that attack will occur. There have been numerous “surprise” physical attacks in history, but typically there is an observable buildup of supplies and forces.

2.4 Specialist Workshop

Our third goal was to assess and validate the current state of the art in the academic, defense, and other communities through a focused technical workshop at the NU level. The IST-129 RTG organized a Specialist Meeting entitled “IST-145: Predictive Analytics and Analysis in the Cyber Domain” in Subiu, Romania, in conjunction with the 40th IST Panel Meeting October 10–11, 2017. The purpose of this Specialist Meeting was twofold: first, to look at the science of predictive analytics in general and secondly, to implementations of predictive analysis, specifically with regard to predicting adversarial cyber operations. This Specialist Meeting identified several areas where researching prediction both within and outside the cyber domain is taking place. While some work has been done, in the opinion of the committee, not enough; much work still needs to be done in both research and implementation. Results from this Specialist Meeting are detailed in Section 10 and are organized into five areas: key results and findings as identified by the committee, some general observations on the practice of prediction, and then some recommendations for the cyber modeling, cyber analytic/algorithm, and cyber prediction communities. Furthermore, the proceedings of the Specialist Meeting were published as a DEVCOM Army Research Laboratory report (McCallam et al. 2019).

2.5 Roadmap and Next Steps

Our final goal was to develop an initial roadmap for development of a comprehensive set of methodologies, technologies, and tools for advancing the proactive Predictive Analysis of Adversarial Cyber Operations. The committee completed this goal through the recommendation of some TAPs and specific research concepts for follow-on activities. The committee noted that in several areas there was a lack of relevant research and the documented work, findings, and recommendations of this RTG represent material available to researchers for further development of predictive analysis in the cyber domain.

3. Key Initial Assumptions

At the initial meeting of the RTG, we made foundational assumptions to help guide and focus the work of the committee. This included classification of the committee including the information to be studied, other domains that could be considered relational or adjacent, modeling and simulation (M&S), and some observations on COAs.

One of the challenges faced by the RTG was the information that would be required for the RTG to effectively perform its mission. We facilitated that discussion and bridged the gap between military and nonmilitary domains since any solution or way forward in solving this problem would work in both domains. Secondly, we decided that all information used within the committee would be unclassified. While there was initially some discussion about the need for classified information, the RTG quickly discovered there was little information at all in this area at any classification level that would directly contribute to a solution path. We were not interested in *how* any information was gathered; instead, we were interested in *what* that information was. With a lack of information in general, we elected to reach into any domain beyond just the military one to utilize any findings or best practices being used. The committee decided that all information used and reported by IST-129 will be sourced and validated as open-source material. We believe that keeping the work of the RTG at the unclassified level will facilitate the ability to share information in the future. This could have a beneficial impact on the speed at which a solution is developed. This assumption was carried throughout the entirety of the RTG's existence.

Secondly, we also believed that noncyber areas had developed approaches to analytics that might be useful in cyber-domain applications. For example, many commercial companies use deep learning and other forms of analytics for inventory control and maintenance prediction. The assumption was that prediction approaches can be easily transferred from one domain to another. What we discovered while performing our research was the assumption was not entirely accurate and that was due more to the lack of knowledge of what data would be necessary for prediction (and conversely, what data is *not*), along with what the constructs of that data would look like.

We also observed that some predictions in the commercial domain were directly dependent on previous events and felt this could be relevant in the research. For example, in the commercial problem of what merchandise to keep on hand and then how to display it, the prediction is focused on specifics, whereas the previous observations are directly related to the future view. As the research unfolded, we realized that in the cyber domain this relationship may not exist. Our conclusion was that while an individual cyber event *could* be related to a previous event, it did not follow that subsequent cyber events were directly related to previous ones.

One area of initial interest was M&S, which besides being an established and well-defined area, may offer constructs for evaluation of solutions. Additionally, since many nonmilitary areas deploy machine learning (ML) and data mining (for inventory control, maintenance prediction, etc.), there was the potential to have components from established areas provide some level of guidance for the RTG. We quickly observed most M&S activities were based on richness of available and realistic data. The cyber domain has long suffered from a lack of data that can evaluate solutions. So while the M&S domain offers promise in the future to evaluate potential solutions, this will require different data sets that can represent threats using known vulnerabilities with unknown exploits.

We also discussed COAs at the initial meeting of the RTG as a potential area for investigation. Although we noted that COAs could be expressed as variable actions or binary (yes/no) outcomes, we dropped the investigation of COAs from going further.

Overall, we maintained the requirement for unclassified data and used the observations we made on M&S and other areas of prediction to guide the program of research.

4. Predictive Analytics Foundational Discussions

4.1 Challenge of Data and Data Analytics in the Cyber Domain

Data analytics can be used to answer different sorts of questions relating to the occurrence of events:

1) What happened? — *Descriptive Analytics*

2) How did it happen? — *Diagnostic Analytics*

3) What will happen next? — *Predictive Analytics*

4) What should I do? —– *Prescriptive Analytics*

In the cyber domain, each question poses challenging problems, not least due to collection and management of the required data.

For the first two questions, to understand what happened in an attack and conduct forensics afterward, we need to collect trustworthy records and logs that cannot be tampered with by an attacker. Currently, data collection happens to a good extent at (some) nation-state levels and within cyber-capable enterprises—often via managed security services provided by security technology companies who collaborate on threat intelligence gathering using Security Information and Event Monitoring (SIEM) systems connecting to security operation centers. But in many countries, smaller businesses, public sector organizations, consumer networks, and even critical national infrastructure networks have less-adequate cybersecurity controls and data collection and are all at risk of being targets in cyber warfare.

The second two analytics questions, what will happen next and what actions should be taken, are more challenging still. To make good future predictions, we need comprehensive data sets containing histories of previous events and, ideally, knowing the ground truth for the signals in the data that preceded them. Even with such knowledge of past attacks, future attacks are designed by our adversaries to look different than previous attacks. So, predicting the movement of previously seen, known attack techniques will be possible, but predicting the appearance and format of new "unknown unknowns" from zero-day vulnerabilities or entirely new attack vectors is a problem not tackled by standard methods in ML. However, here we suggest that integrating analytics with threat modeling offers a way forward.

4.2 Data for Cyber Analytics

In the cyber domain, the main sources of data include the following:

- Network traffic of various kinds captured by devices on the network (Transmission Control Protocol [TCP] packets and flows, User Datagram Protocol [UDP] counts, Domain Name System [DNS] queries, Border Gateway Protocol [BGP] messages, etc.);
- Log files on endpoints and intermediate servers capturing server and application-level actions (web servers, mail servers, application-level gateways, firewalls, intrusion detection systems, etc.);
- And sometimes, behavioral information capturing the interactions of humans with systems, as well as other organizational-level data and threat intelligence categorizing known attacks, attack vectors, and propagation methods.

As well as data itself, certain metadata is useful. Metadata may include traffic communication patterns and timing and latency information. This is becoming increasingly useful with the rise of end-to-end encryption and the need to "see inside" closed network systems. The survey of data-driven methods for intrusion detection by Buczak and Guven (2016) gives more detail of data types and other useful features.

4.3 Concept Drift and Adversarial Influence

Recognizing malicious network activity is not like recognizing pictures of cats—attackers continually develop new malware and exploit newly discovered vulnerabilities with the hope of evading current detection mechanisms. This is a prominent example of the problem of *concept drift*, where statistical properties of the target variable change over time. A model that is trained at one point in time becomes less accurate over time and needs to be retrained.

Whether detecting attacks or predicting future trends, ML, and other artificial intelligence (AI) systems are at risk of adversarial influence. There are two attack methods:

- So-called *poisoning attacks* that occur during the training phase (which will be continuous in unsupervised, semi-supervised, or retraining). These are possible if an adversary can provide untrusted input to the algorithm, skewing the trained model. For example, an attacker can hide a low-rate data exfiltration among ordinary traffic by seeding the model to expect certain kinds of connections.
- Specially crafted *malicious counterexamples*. These have been demonstrated strikingly for image recognition problems since 2014 and for many ML algorithms and applications since, spawning research in constructing families of such malicious counterexamples and designing systems that are robust against them. In the cyber domain, such malicious examples could be used to trigger false alarms in detection systems or fool predictive analysis into making false predictions.

Alongside these strategies to corrupt detection systems, the usual strategy of an attacker is of course to evade attacks by obfuscating and diversifying malware to avoid signature-based detection, concealing payloads, and so on.

4.4 ML Methods

Machine learning is a collection of statistical and heuristic methods in AI used for modern data analytics, such as classification problems, data clustering, or predicting outcomes. The basic paradigms are detailed in Table 4.

Table 4 Basic ML paradigms

Paradigm	Example problem in the cyber domain
Supervised learning	Distinguishing malware from benign software (Kolter and Maloof 2006)
Unsupervised learning	Clustering malware into similar behaviors (Perdisci, Lee, and Feamster 2010)
Semi-supervised learning	Evolving a behavioral model of network normality (Ashfaq et al. 2017)
Reinforcement learning	Guiding a software defense bot in a hostile environment (Zhu, Hu, and Liu 2014)

Each paradigm has a set of well-studied algorithms and techniques associated. The choice of best algorithm for a problem domain depends on several factors, such as the size of the data set or the number of training features and their kind.

Recently, notable advances in ML have been made with "multi-algorithmic" *ensemble* methods (boosting in the supervised setting, which improves accuracy and reduces bias) and with *deep learning* methods (such as recurrent neural networks, which are able to capture temporal structure in data inputs).

By now, the scientific literature contains hundreds of experiments with different ML algorithms with cybersecurity applications. However, most of the current research concerns *detection* methods, used in Problems 1 and 2 mentioned previously in Section 4.1, rather than actionable predictive or prescriptive analytics.

4.5 Questions for Predictive Analytics

Some examples of data-driven predictions we would like to make in the cyber domain are the following:

1) When will the next cyber-attack on a network or system occur?
    - How long do we have to prepare; can we install defenses in time?
2) What kind of attack is most likely to occur next (or be in progress) on a given network or target endpoint?
    - For example, will it be a denial of service, an insider attack, an advanced persistent threat (APT) becoming active, or a virus or worm?
3) Where is the next attack most likely to come from?
    - What will be the most probable attack vector?
    - Who is the most likely threat agent? (future attack attribution)

4) What is the likely impact and cost of the next attack?

    - What assets will be at risk?
    - How much downtime or network interference do we need to endure?

The current prevalent types of data collected for cybersecurity intrusion and attack detection mentioned previously are mainly concerned with operational procedure and immediate security responses (if my network neighbors see a worm attack, I should defend against it). Knowledge of vulnerabilities and patches, malicious network traffic observations, and prevalent malware are all useful to take preemptive responses such as urgently updating an application suite or reconfiguring firewall rules. This kind of data may also be used to address the first two kinds of prediction. To go further with prediction, we need more forms of data. For example:

- **Threat intelligence data**. Particularly data used for *strategic* intelligence, which consists of values put on core assets as well as catalogues of known or anticipated threat actors and their TTPs. Existing open-source or commercial *operational* threat intelligence data feeds could also be used in the future to help build predictive models of trends in vulnerability types, malware variants, or suspicious Internet Protocol (IP) addresses (which may be helpful to track hacker groups).
- **Historical attack data**. To predict how future attacks will propagate and have impact, we need historical data for previous attacks and especially their cost. Unfortunately, this can be scarce because organizations may not have seen a wide range of attacks, data has not been kept, and it may not be shared outside organizations (in commercial or government civil or defense sectors).

Cyber incident recording and reporting has had a varied history. Commercial organizations in many countries are now legally obliged to report cyber breaches and be fined for them (especially when personal data may be involved, such as in Europe under requirements of The General Data Protection Regulation). National Computer Security Incident Response Teams (CSIRTs) have evolved to collect and share data on cybersecurity incidents (primarily or wholly at a civil level), engaging with local and global law enforcement, as needed for coordinated takedowns of globally operating cyber-criminal gangs. The United Nations recommends CSIRTs be operated by each country, although their format varies and there is a debate about the effectiveness of this (Skierka et al. 2015).

Some independent organizations, for example, the not-for-profit Shadowserver (Shadowserver Foundation 2019), attempt to collect data worldwide on current malicious internet campaigns and coordinate with large infrastructure companies and internet service providers as well as law enforcement.

Despite these rich data sources, research on effective predictive analytic techniques is impeded by a lack of openly available data sets that contain organization- and asset-specific impact and severity information, as well as information on sources of attacks. (The Common Vulnerability Scoring System scores recorded by the National Institute of Standards and Technology [NIST] against common vulnerabilities and exposures in the US National Vulnerability Database do not provide a useful level of granularity nor the ability to transfer results between organizations.)

Openly sharing this kind of information is likely to be difficult, although emerging privacy-enhancing technologies, such as those being explored in the US Department of Homeland Security (DHS) Information Marketplace for Policy and Analysis of Cyber-risk & Trust (IMPACT) program's Framework for Information Disclosure with Ethical Security (FIDES) project may offer a future way forward (Galois 2019).

4.6 Evaluating Results and Comparing Tools and Methods

In the early stages of new technology product types entering the market, there can be a lack of established benchmarks or performance criteria so customers do not have good ways to discriminate between products. It can even be in vendors' commercial interest to prohibit comparative study. This leads to *attribute substitution*, where customers rely instead on heuristics to make complex product choices, such as brand reputation. This is the present state of the market in several cybersecurity product types, especially ones incorporating AI methods.

In software assurance, the NIST Software Assurance Metrics and Tool Evaluation (SAMATE) effort has been trying to introduce objective criteria for measuring tool products in software assurance (NIST 2018). We hope that future marketplaces for data-driven cyber analytics and AI will also be subjected to more uniform evaluation and comparison. To do this will require establishing agreed-upon benchmark testing data sets, which is itself a challenge for reasons mentioned previously.

Also, partly because of a lack of suitable data, scientific evaluation of data-driven methods in cybersecurity has had drawbacks. For example, with fixed data sets it is possible to refine algorithms continually to improve key metrics such as precision and recall, reaching in the high 90 percentiles, but perhaps *overfitting* to the training data. A recent paper (Pendlebury et al. 2019) points out further two kinds of bias seen in some cybersecurity ML research work:

1) *Spatial bias,* in which training and testing data are not drawn from the same distribution as the actual target data, for example, training anomaly detection methods on data sets where there is an artificially high amount of malicious traffic (some researchers have used 50/50 splits whereas realistic data has under 1% of malicious traffic in the long run).

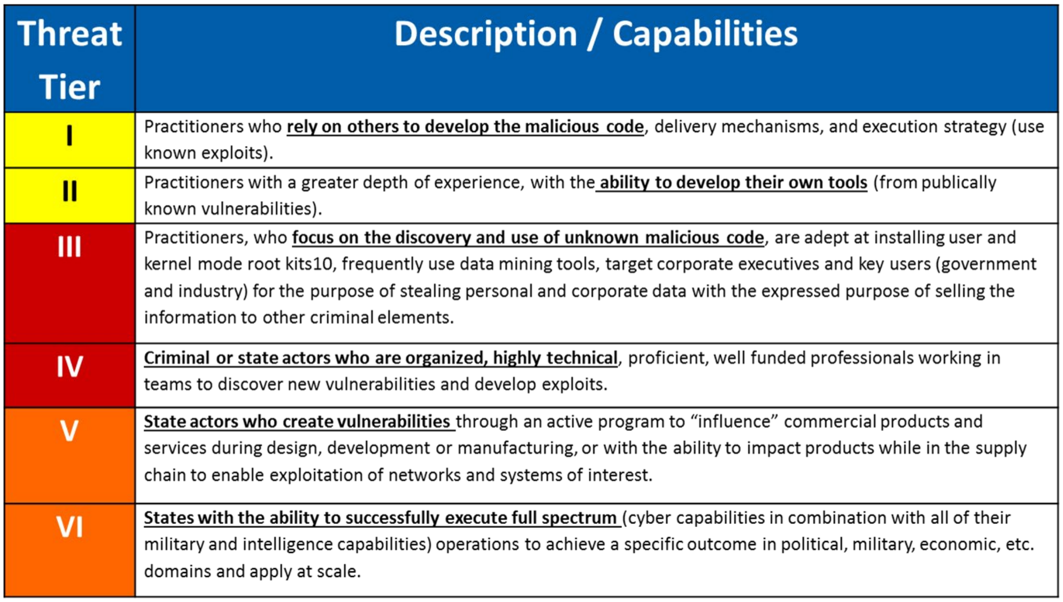
2) *Temporal bias,* in which methods are trained given "future knowledge", for example, by evaluating detectors against old data (known attacks) rather than assessing their performance against newer data (attacks not seen during training); of course, methods should ideally be *robust* and adapt well as the adversary evolves.

Emerging predictive analytic methods will need to be evaluated carefully bearing in mind these kinds of bias.

5. Discussion of Threats in Terms of Capability

Initially, the committee selected a definition of the threat (Table 5) as found in the DSB report that defined cyber threats in terms of capabilities as opposed to identifying specific groups (Gosler and Von Thaer 2013). This allowed the work of the committee to address threats in terms of capabilities, which is universal in terms of the cyber threat but avoids potential classification issues of specific group identification. This capability description has six levels organized into three bands of capabilities. Levels I and II concentrate on threats that leverage known vulnerabilities using known exploits. Levels III and IV concentrate on threats that focus on known vulnerabilities using unknown exploits. Levels V and VI are more the state actors that have the capabilities to create unknown vulnerabilities and associated unknown exploits. From a financial investment point of view, operating at Levels I and II is very cheap. The investment in capability development escalates with Levels V and VI and is very expensive. From a focus area, the committee eliminated Levels I and II since these are deterministic areas that are addressed through signature detection. The committee elected not to "boil the ocean", so decided to focus the activities on Level III, referred to as the (known, unknown) representing threats that focus on known vulnerabilities using unknown exploits.

Table 5 Description of cyber threats with respect to their capabilities



Threat Tier	Description / Capabilities
I	Practitioners who **rely on others to develop the malicious code**, delivery mechanisms, and execution strategy (use known exploits).
II	Practitioners with a greater depth of experience, with the **ability to develop their own tools** (from publically known vulnerabilities).
III	Practitioners, who **focus on the discovery and use of unknown malicious code**, are adept at installing user and kernel mode root kits10, frequently use data mining tools, target corporate executives and key users (government and industry) for the purpose of stealing personal and corporate data with the expressed purpose of selling the information to other criminal elements.
IV	**Criminal or state actors who are organized, highly technical**, proficient, well funded professionals working in teams to discover new vulnerabilities and develop exploits.
V	**State actors who create vulnerabilities** through an active program to "influence" commercial products and services during design, development or manufacturing, or with the ability to impact products while in the supply chain to enable exploitation of networks and systems of interest.
VI	**States with the ability to successfully execute full spectrum** (cyber capabilities in combination with all of their military and intelligence capabilities) operations to achieve a specific outcome in political, military, economic, etc. domains and apply at scale.

Note: Derived from p. 22 of the DSB report (Gosler and Von Thaer 2013).

Subsequently, an updated approach to categorizing threat capability was published in a GAO report (GAO 2018) that provided a more refined approach to defining cyber-adversary capabilities. This structure is described in Table 6.

Table 6 Updated definition of cyber threats based on capability (adapted from Lamolinara 2018)

Actor capability tier level	Relative capability	Details
Level 1	Nascent	Predominantly composed of minor nonstate actors and poorly organized/resourced state actors with little or no organized cyber resources. These actors primarily exploit known vulnerabilities and use readily available tools/methods, although have some limited ability to create tools. Targeting is specific, mostly local targets for personal, financial, and strategic gain. They possess little or no knowledge of a network's underlying systems or industry beyond publically available information.
Level 2	Limited	Composed of state and nonstate actors who are able to identify and target for espionage and attack readily accessible unencrypted networks running common operating systems using publicly available tools. Actors at this level may be well organized and resourced and can be determined/persistent. They are also capable of discovering new vulnerabilities.
Level 3	Moderate	Composed of actors who are able to use customized malware with better operations security to conduct wider-range intelligence collection. Moderately capable state and nonstate actors who are highly organized and deeply resourced with the ability to gain access to more isolated networks and create short-duration effects against critical infrastructure. These actors can create new zero-day attacks and customize malware.
Level 4	Advanced	Highly capable state actors that have the capacity to conduct complex, long-term cyber-attack operations combining multiple intelligence disciplines to obtain access to high-value networks. They can exploit and affect supply chain and develop advanced attacks. For attacks from this threat level, leadership understands the risks and consequences.

How We Limited the Scope of the Research

Predicting adversarial behavior in cyberspace is a large domain and we needed to limit the scope of the research to ensure the RTG could not only complete its research work, but also to give a starting point. We elected to restrict the threat domain to only those threats that use known vulnerabilities with unknown exploits. This aligned with current literature that addresses threats less by geography and organization and more by overt capability.

6. Key Discussions and Influencing Factors

6.1 Temporal Issues/Laws of Physics

Two of the major challenges in predictive analytics and cyber defense are the speed at which decisions must sometimes be made, and the lack of observable physical phenomena.

Speed in cyber requires efficiency in analytics/analysis, which is contrary to current state of the practice. Autonomous systems must often act as quickly as possible, with time for only rudimentary analysis. A truly autonomous and predictive system would need to consider both short-term predictions (for example, in tactical situations), but also perform longer-term trend analysis (for example, to increase supplies or shuffle logistical resources). All of these predictions would need to be made quickly enough to still have enough time left for the recommended actions to occur and matter.

In other fields, such as aerospace, the path that a missile takes can be easily predicted (or projected) based on its position and velocity. In orbital mechanics, the six Keplerian elements can completely describe (or predict) an object's orbit. The objects must obey well-known laws of physics. Cyberspace has fewer physical constraints; therefore, the realm of possibilities is much larger and prediction is much more difficult.

The laws of physics still apply to the wires and circuit boards that compose the computers of the cyber world. However, the same laws of physics that govern objects in the physical world do not apply in the same way to the virtual world of cyber attackers and defenders. In contrast to the aerospace characteristics of a missile, a packet traversing a network is not limited by physical distance, cannot control the route it takes, and might arrive at a target many different ways from many different directions. This difference in physical constraints makes prediction much more difficult in the cyber realm.

Before a system can do prediction, it must first have some sort of environmental knowledge, and then some sort of detection capability. Environmental knowledge is difficult because of factors like APTs, unknown vulnerabilities, zero-day attacks, and spoofed traffic. Detection is difficult because of these same factors, as well as the volume of network traffic or log data that is legitimate and must be analyzed, filtered, or discarded to find the "needle in the haystack" of a discrete cyber-attack. If a system is mostly secure, and has good environmental knowledge and an accurate detection mechanism, only then can it begin making accurate predictions, and possibly recommending COAs.

Ironically, the paradox of prediction still applies in the cyber realm. The paradox of prediction (DOD 2013) occurs when Group B is preparing to attack Group A. Group A somehow gains intelligence or detects the preparations for attack. Group A increases its defenses. Group B cancels its planned attack because of the increased defenses. Because the attack from Group B never occurs, Group A then questions the accuracy of its predictive capabilities. This still occurs in cyberspace, both to the humans monitoring cyber systems and to AI systems trying to operate autonomously. This type of proactive determination (by Group A) makes measuring the accuracy of the initial prediction extremely difficult.

These factors all contribute to the difficulties of doing predictive analytics in a timely and accurate matter. Each of the component problems is well known and are active areas of research and development. Our research found some systems trying to do predictive analytics within limited scopes (e.g., attack graphs leading to vulnerability exploitation). However, the field is not very mature and much more research is needed.

6.2 Discussion of TTPs

Dominance in cyber conflict and cyber-adversary dynamics will depend on our ability to recognize and eventually anticipate changes in adversaries' TTPs in cyber operations. TTPs are profiling certain threat factors and we try to prove that in the future, commanders on battlefields could have automated systems to recognize adversary tactics and tools. Basic TTPs are for the concept of "how" and "what" of adversary behavior. Some TTPs are using key factors for information technology (IT) infrastructure, victim targeting, attack patterns, and different malware variants. Characterization of malware or malware variants is one of the basic TTPs (like Zeus, Conficker, special variant of Stuxnet, etc.)

To achieve certain categorical results in cyber prediction, we have to distinguish dual use of the TTPs for civilian and military use and technically distinguished mobile and organizational (enterprise). TTPs for air, maritime, and land are written in different standards, but areas of cybersecurity especially focused on prediction are not scientifically explained and standardized yet. We found that MITRE classification* could be used for unknown/known threats but category postattack and prediction is missing (Fig. 1). Methodology for MITRE is a synthetically/

*MITRE in 2018 recognizes cyber matrices as PRE-ATT&CK, enterprise for different platforms and mobile. The same categorization is being used for Tactics and Tools. Used for cyber prediction, could be Priority definition planning, Priority definition direction, and three levels of information gathering (technical, people, and organizational). In general, they use 15 different areas and it is TTP breakdown for computer security.

unrealistic solution because it depends on historical events, covers only a portion of all adversary activity that exists, and has some physical limits like time constraints without "user noise".

Another assumption we found is about verifying offensive capabilities with different known or unknowns. TTPs are getting more complex in the way that an adversary could use offensive tools' cyber capabilities portfolio. In that portfolio, an adversary can influence in gaining and maintaining situational awareness of military units, acting against offensive and defensive infrastructure.

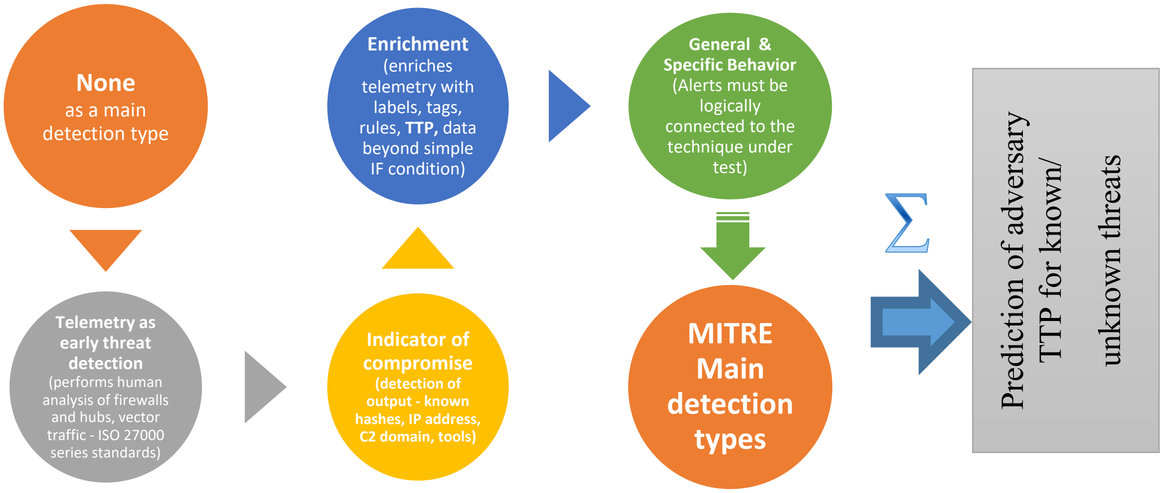


Fig. 1 MITRE logical flow evaluation techniques includes TTP source: https://attackevals.mitre-engenuity.org/APT3/detection-categories.html

We studied 12 papers on cyber profiling via TTP and also have some updated e-resources connected in a sense of cyber prediction.

Developing new innovative "hack-proof" TTP methods of command and control is necessary and we need to analytically predict what kind of threats are coming. More sophisticated cyber-concentrated capabilities with automated TTPs could develop our asymmetric advantage in cyber-battlefields. Multisensors[†] are needed to be integrated in TTPs to prevent and predict future "bad actors".

No specific software or hardware tool for prediction analysis is exact and we can always predict with wrong conclusions, especially in *software failure*. Adversary unpredictability will induce uncertainty or invalidation in security of military computer systems. Error is a component for prediction uncertainty. Measurement

[†] Mobile e-camouflage with thermal, infrared, or radar full-spectrum sensors.

theories are not proven yet on predictive analytics to quantify errors like calibration against known security value, or statistical comparison to use big malware data.

We uncovered several challenges in basing the predictions and the analysis only on TTPs:

1) Postdetection researching capabilities are necessary when a threat bypasses defenses or uses new facts to enter a network to minimize damage.
2) TTPs need some unique categorization (low-, middle-, and high-level tools and tactics).
3) Biological approach with biological viruses and how to predict new threats could be very helpful.
4) Tactics are divided in general into 10 topics (Persistence, Privilege Escalation, Defense Evasion, and so on…)
5) Prediction TTPs should be systematically accessible 24/7.

6.3 Handling of Attacks by Layered Structures

Prof Peeter Lorents, member of the IST-129 RTG, provided a lecture on handling cyber-attacks through layering structures. The structure of attacks can be considered as a multilayered structure (e.g., a layer of triggered effects, a layer of means of generating effects, a layer of capability needed to use resources, and so on.). Attacks can be seen as "paths" that run through different layers in a layered structure. To obtain initial robust estimates of the possibility of attacks, one option is to use the so-called layered structure with relationships between the elements in the layers. Some examples of offense layers were defined in terms of a complex 6-tuple consisting of the following:

- The *resulting situation*, denoted "Res", defines the desired strategic result of the attacked in terms of overall impact to an adversary. There could be multiple resulting situations or processes (Res1, Res2, …).
- There are also *inducers*, denoted "Ind", that define the resulting more tactical situations or processes that would include effects, events, developments, and so on. (Ind1, Ind2, …).
- Each inducer, Ind_i, has an *option* ("Opt") that indicates the target of the inducer. Some of these options could include equipment, software, systems, and so on. (Opt1, Opt2, …). Note that this also identifies a vulnerability attack point.

- Each adversary has innate *capabilities*, denoted "Cap", that are the skill or knowledge sets required to realize the possibilities of implementing inducers. This is a similar approach defined by both the DSB report (Gosler and Von Thaer 2013) in Table 1 and GAO report (GAO 2018) in Table 2 to communicate threat capability level. This major difference is that Cap defines specific capabilities that could be attributed to an adversary.
- In an effort to more specifically identify an adversary, *Hol* is defined as the holders of specific capabilities. This could use varying forms of identification/attribution such as groups of people, communities, affiliations, nationalities, and so on.
- Finally, the stimulus or motivation of the adversary is denoted *Sti*. Stimuli and motivations could be economic, political, nuisance, and so on.

During the discussion, several observations were made. One is that these 6-tuples are compound definitional structures. For example, specific holders (or threats), H_i could be motivated by several stimuli (Sti_j, Sti_k, Sti_{k+1}), creating a definition about the various motivations of that specific holder. Similarly (and more realistically), holders could possess multiple capabilities that each utilize multiple options targeting specific options. The structure allows for developing recursive holder definitions—very useful for getting specific about a specific adversary's complete capability suite.

Given this complex set of recursive relationships, it is possible to evaluate the chains of existence (e.g., the probability that *some adversary X has the capability Y* and *the capability Y is necessary for X to be able to create certain malware Z*). How this may work in the prediction domain could be assessed using similarity approaches. For example, if we are looking for an adversary to have the capability, Cap_x, we could examine other adversary relationships and gauge the similarity against holders of Cap_x.

The discussion continued providing conjecture at ways this approach could be used, all of which would require some level of rigor to prove usefulness. It is possible attacks can be presented as a particular conjunction, where the conjuncts (or conjunctions operands) are claims that the elements from the layers are in some relevant relationship. Using estimates and procedures for assigning them, you can also assess how plausible it is to have one or another potential attack.

Such an approach is, to a certain extent, even suitable for "crafting"—if someone is able to define the appropriate layers, the things they contain, the interrelationship between things, and the judgment about it. True, such a "craft" may not produce a very accurate result, but it is relatively quick and can be done with little effort. This may provide an indicator if the conjecture or relationship is plausible, possibly plausible, absolutely not plausible, and so on. In the case of structural handling of attacks, methods of structural similarity/specificity detection can be applied to better predict the occurrence and possible outcomes of any attacks with a structure.

6.4 Similarity of Situations and Processes

Prof Peeter Lorents, member of the IST-129 RTG, provided an additional lecture on using similarity as a means to address prediction. In decision support systems and autonomous decision-making systems, one way is to model human decision-making and in particular, the human "techniques" of reliance on similarity. The reasoning in the discussion is the human tendency to try and relate a new situation to one that is "similar" in the past. The discussion focused on the similarity of descriptions of situations and developments with the aim of using similarity of known things to help us understand the similarity of unknown things. This discussion is aimed at applying this concept to situations and developments in cyberspace: the computer systems, networks, and so on. While finer points of the discussion and examples did not pertain directly to adversarial prediction in cyberspace, the overall concept did appear to have promise.

In mathematical terms, similarity looks at the intersection over the union of multiple data sets. In such a case, it is possible to calculate numerical estimates that can be used to decide how credible a potential attack is and what further development is. For the evaluation of similarity, it is possible to use an approach whose roots originate with the Jaccard coefficient, a value ranging from 0.0 to 1.0 that measures relative similarity between finite sets. While some of the relevant mathematical tools are not too sophisticated or complex, they have certain limitations. For example, the Jaccard coefficient is not as accurate with small data sets but has increasing accuracy over larger data sets. Given that, it is still possible to create appropriate algorithms and IT-based analytic solutions to compare new cyber information (Is A_n an attack?) with known information (B_n was an attack) and compute a similarity coefficient.

Decisions leading to responses to cyber-attacks need to be made very quickly. One important step in shaping appropriate decisions is to anticipate the immediate situation and be in a position to respond to future developments. An important feature when dealing with cyber-attacks is we tend to have large data sets (a condition favorable to using the Jaccard coefficient) and the requirement that high processing speeds are necessary to analyze data and form decisions.

An observation was made that reliance on a similarity/nonsimilarity assessment is more suited to automation since the relevant parameters (speeds and volumes) are better machine than human satisfied. The point of similarity is to be in a position to choose an option that has succeeded in similar situations and, just as important, avoid doing things proven to be unsuccessful in similar situations. Consequently, well-founded and automated methods and tools would be needed to assess similarity/nonsimilarity and be able to apply it to adversarial behavior prediction.

Similarity may have some impact into being able to predict behavior, or at a minimum could be used to select high potential courses of action through high Jaccard values across known data sets.

6.5 STIX

STIX is a language and structured format used to exchange cyber threat intelligence (CTI). It is open source and widely used. The STIX framework intends to convey the full range of potential cyber-threat data elements and strives to be as expressive, flexible, extensible, automatable, and human-readable as possible (Barnum 2012).

When the RTG was conducting a literature survey on relevant, related research (discussed in Section 9), it quickly became apparent we needed a method to categorize the various survey papers. The survey papers we considered often made different assumptions about the environments in which they were operating and what information was available. The papers were difficult to compare because they were usually making different types of predictions, or predicting different types of events.

To more easily compare the papers and discuss the various types of prediction, we needed a framework. Members of the RTG familiar with CTI noted similarities with STIX. STIX is a framework used to convey CTI. If a predictive analytic system is able to predict and/or detect a threat, it must still communicate that information in a useful way. So comparing predictive analytic systems based on what categories within STIX they used as inputs and outputs became self-evident. The RTG quickly adopted this as a method to use for categorizing the various papers and the types of prediction being made.

Even with this framework, accurate comparisons of predictive techniques are difficult. Most of the papers predated the framework and were difficult to categorize, or used terms that could fit into multiple categories. Most papers assumed different types of data as input and provided different types of data/predictions as output. Section 9 describes how we categorized the survey papers found and what types of information we tried to extract.

For simplicity, the RTG used STIX 1.2. However, we are aware of STIX 2.0, TAXII, and CybOX. The proliferation and wide adoption of these frameworks speaks to the need for, and success of, these frameworks. Cyber threats are growing in number and severity. Cyber-defense systems must grow and expand also. As the number of cyber-defense systems grow and become more complex, they will still need a framework like STIX to communicate efficiently. These frameworks are also useful for sharing information about threats while predictive techniques are being developed.

6.6 Examination of Boyd's OODA Loop

There exists many loops for what adversaries do with controls and products. Tactical/operational-level decision-making with Boyd's OODA loop for military organizations is fraught with difficulties in the cybersecurity area (Boyd 1986). A measure of prediction with some mode of behavior with higher probability rate is useful. Certainty with prediction in OODA loop theory could quantify precision of physics elements in cybersecurity. Combat engagements in cyber battlefield needs systematic analysis of known TTPs and gives understanding of what is going to happen. In cybersecurity, long- and mid-term predictions are worthless; we need short-term or even almost-current predictions.

At the third meeting of the committee, we invited Wing Commander John McCarthy of the Royal Air Force to initially discuss with him both visualization and decision cycles as a potential way to address the predictability problem. John Boyd's OODA—Observe, Orient, Decide, Act—loop is a well-known model for behavior in many domains. Fundamentally, it defines the decision and action cycle of any organism and has been used to model human decision-making (Angerman 2004). The group discussion elected to examine how/if this would apply in some way to the problem of predictive analysis. The group did a short academic paper search in the OODA loop as it applied to cybersecurity and found some information, but for the goal of predictability it came up short.

We did find several variations in terms of images of a cyber OODA loop and they are shown in Figs. 2–4. The interesting thing about the various approaches was the characterization of the orientation phase. Boyd's original loop had five "influencers" for the orientation phase: cultural traditions, genetic heritage, analysis and synthesis, new information, and previous experience. The group focused on those influencers and how a similar framework could be developed for cybersecurity. In the orientation phase of the OODA loop, we discussed how culture and genetics are represented in the cyber world. We had a parallel discussion on sports, in particular football (soccer), where learning how to play a particular style (Dutch total soccer, Brazilian ginga, etc.) has a high influence on how a player approaches the game. We believe that this is also the case with offensive cyber. The way an adversary learns the craft will influence how they execute cyber operations. This led us to equate the cultural heritage/traditions in Boyd's concept to craft learning in ours. The second discussion area was for genetic heritage. Genetic heritage in the real world relates a person's cellular/organic makeup. We equated organic makeup of cyber events to event DNA. Notionally, this made sense. The result of the analysis of other work and our discussions resulted in Fig. 5, which better explains the impacts of learning adversarial cyber operations on where to focus analytic attention.

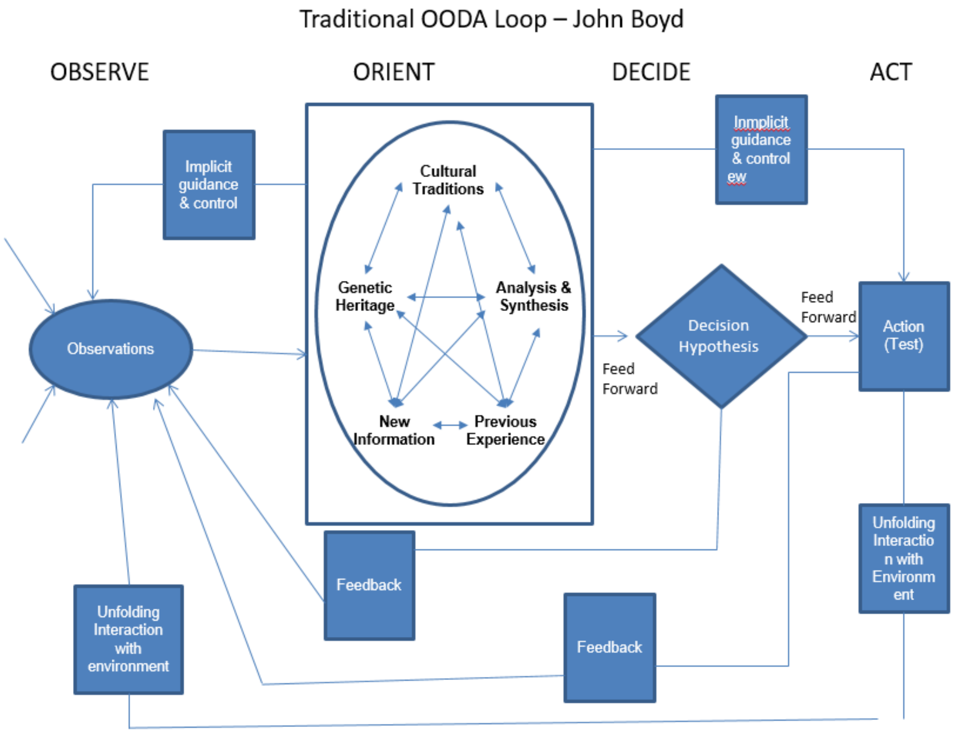


Fig. 2 Boyd's original OODA loop (Boyd 1986)

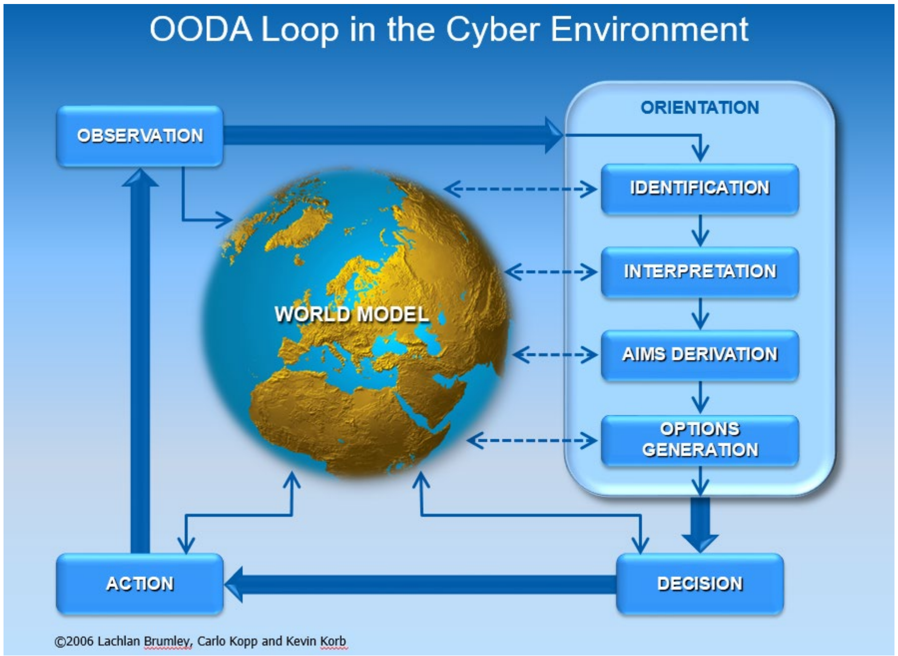


Fig. 3 Good representation of the orientation step of the OODA loop (adapted from Brumley et al. 2006)

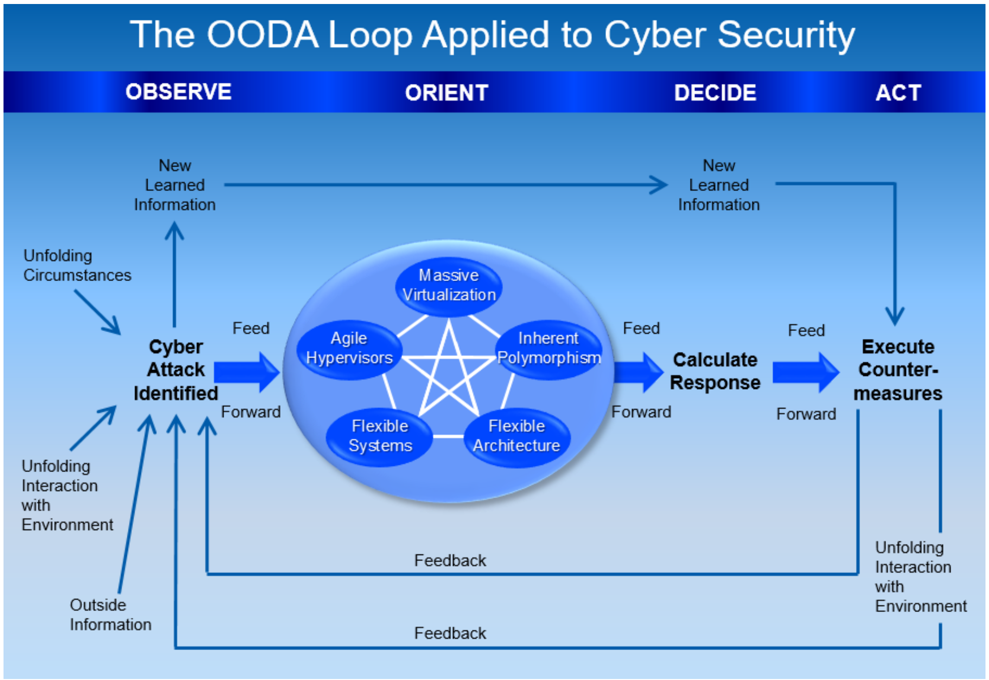


Fig. 4 Unsourced OODA representation with interesting orientation approach

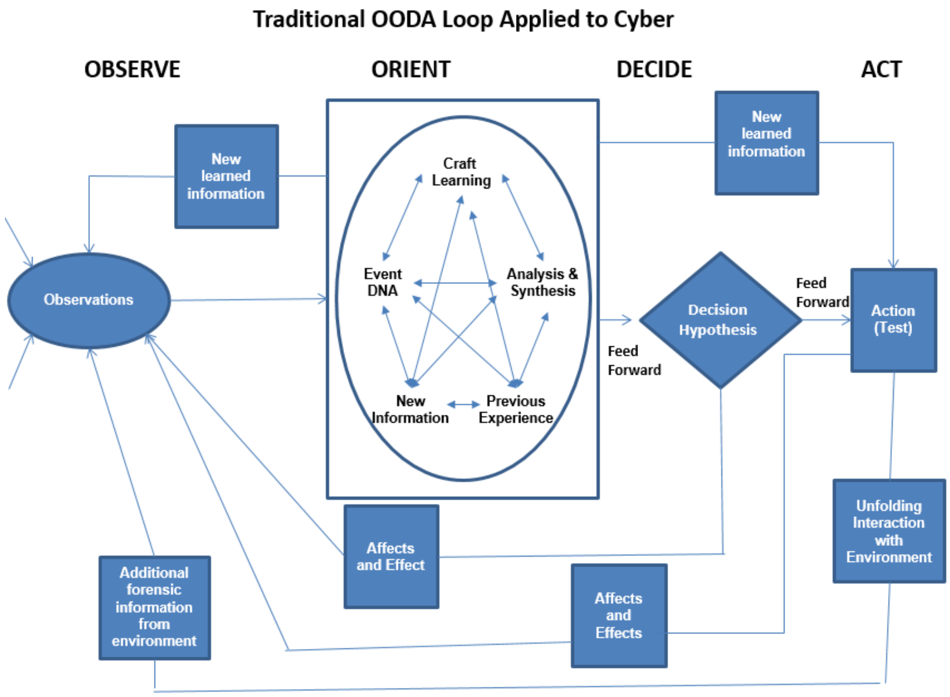


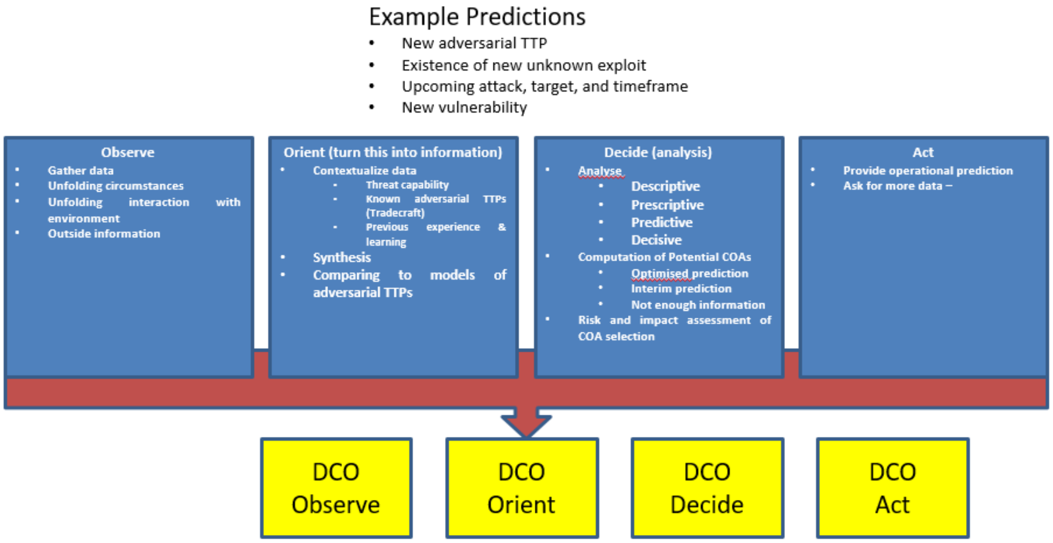
Fig. 5 IST-129 postulation of OODA loop with adaptive orientation phase

The notion of adversaries using established TTPs was surfaced. Several points were made including the following.

If it appears an adversary is using a known set of TTPs, then it was postulated it could be possible to predict the next event. This was disputed with a lottery argument. Some lotteries show how many times in the last *n* drawings of numbers a particular number was drawn. The implication, although false, is that the next drawing of numbers is in some ways related to the previous drawings. Since all the drawings are independent events, there is no previous dependence. This opened up the possibility that an adversary could be using a set of exploits that make up a set of TTPs from Adversary A and then switches to TTPs from Adversary B. This discussion led us to postulate that the problem of prediction in cyberspace *with 100% certainty* is close to impossible.

We then stated two other observations. First, while possibly being unable to predict with certainty, we believe relative certainty can be provided. This could help in better use of resources. For example, the prediction could be made that indicates 90% probability for Event A, 85% for Event B, 35% for Event N, and so on. Knowing this in advance might allow better planning and human capital use. The second observation we made dealt with attack trees. If we could compute all possible attack paths for known vulnerabilities, we could reduce the problem of prediction to one of identification through signature pattern matching.

Figure 6 summarizes our discussion and observations/conclusions for using the OODA loop to predict adversarial behavior.



Example Predictions

- New adversarial TTP
- Existence of new unknown exploit
- Upcoming attack, target, and timeframe
- New vulnerability

Observe

- Gather data
- Unfolding circumstances
- Unfolding interaction with environment
- Outside information

Orient (turn this into information)

- Contextualize data
    - Threat capability
    - Known adversarial TTPs (Tradecraft)
    - Previous experience & learning
- Synthesis
- Comparing to models of adversarial TTPs

Decide (analysis)

- Analyse
    - Descriptive
    - Prescriptive
    - Predictive
    - Decisive
- Computation of Potential COAs
    - Optimised prediction
    - Interim prediction
    - Not enough information
- Risk and impact assessment of COA selection

Act

- Provide operational prediction
- Ask for more data –

Fig. 6 IST-129 discussion on using the entire OODA loop as the orientation step

We discussed two topics: expected predictions and expanding the orientation phase. We also observed that in the decision/action portion of the OODA loop, one possible choice and execution path is to get more clarifying data. The caution here is not to overly make this choice or the consequence is to be caught in an analysis–paralysis cycle. This in essence embeds the entire OODA loop within the orientation phase to expedite decision-making, an observation we considered relevant.

The discussion concluded with some sample predictions that would be useful in predicting adversarial behavior. One outcome could be the discovery of a new adversary TTP that could be a new string of known exploits or the discovery of a new and unknown exploit. The latter prediction could also identify a previously unknown vulnerability. While noting that the optimum prediction would contain what are the components of the predicted attack, what specifically is the target, and when in time this will occur, the realization is this is not within current grasp of technology.

7. Closed Networks and their Direct Implications

There is the possibility that closed national segments of the internet arise in the future. For example, Russia has declared its aim to become “digitally sovereign”. In addition, China, North Korea, and Iran may have their own versions of closed national networks or projects leading to such. One part of “digital sovereignty” is the capability to close off its national segment from the global internet whenever required and maintain operational capabilities of the national segment while doing so (Kukkola et al. 2017, 2019).

Furthermore, it is shown in earlier studies (Kukkola et al. 2017, 2019) that motivation on reaching capability to close national segments of the internet as required may be to achieve a decisive military advantage and to create an asymmetric situation. Moreover, it is proposed that a way to counter the challenge is to proactively learn the properties of the closed networks and to evaluate what their effects are to cyberspace as a whole (Kukkola et al. 2017, 2019). In practice this could mean, for example, to form an ET/RTG to analytically resolve general properties of closed national segments of the internet.

In the discussions, it was concluded that one implication of a closed national network could be that it improved the prediction capability of the specific nation. This observation is based on the potentially improved cyber situation awareness of the nation in question. As a result, there is a chance that a nation introducing a closed national segment of the internet could gain an advantage in the cyber domain specifically related to cyber situation awareness and prediction capability.

However, there are other effects (e.g., related to economy) that may be more significant and it cannot be considered a desired COA for the NATO coalition to contemplate a closed national segment of the internet in any manner. One theme was whether or not NATO alliance should begin researching techniques to construe closed segment networks and studying technology solutions and their vulnerabilities. A potential option is also to organize and update its own policies and possibly propel cybersecurity technologies. One should analyze the effect of some nations’ plans to introduce closed national segments of the internet.

To conclude, it is necessary to understand and analyze the challenge introduced by the formation of closed national segments of the internet. As a first recommendation, we propose to form an ET/RTG to characterize and evaluate the challenge presented by the closed national segments of the internet. The details of the group are given later in this report. As a second recommendation, we suggest modeling the mentioned closed national networks. By introducing the models, one is able to generate more detailed information of the closed national networks. One result of the modeling could be to generate information of which features of the closed national networks could be recreated within open networks. Another result is to uncover which features could not be adapted but should be recognized. As a third recommendation, we advise designing and running experiments to extract and analyze the effects of the closed national networks.

The experiments could be organized based on the models mentioned in the second recommendation and independently as a tabletop exercise (TTX). A TTX could be organized more quickly and it could assist in creating the situation awareness of the implications of the closed national segments of the internet.

8. Invited Industry, Government, and Agency Presentations

During the course of the RTG's meetings, we invited both vendors and government officials to present their views on the topic of prediction and to discuss with us how they saw or were approaching the problem. We had discussions with Avata Intelligence, DHS, Amazon Web Services (AWS), and Extreme Networks. We had hoped to interview a wider user/provider audience, but found only a limited group willing to openly discuss the issues.

8.1 Avata Intelligence: James Pita

James Pita, a cofounder of Avata Intelligence, was invited to present Armorway, which provides Bespoke AI assistance built on a proprietary platform that manages data and implements a large number of algorithms including ML and other methods.

Background. Armorway was developed as a spinout in 2013 from research at the University of Southern California aimed at risk assessment. It was used in a range of projects with the DHS and DOD, starting from federal government projects during the company's research and development stage and achieving some impressive successes in the criminal domain. They widened to target cybersecurity, though this has been a much more challenging domain. (Armorway has since been renamed Avata Intelligence and is now sold to a broader range of market sectors, beyond just security and cybersecurity markets; see https://avataai.com.)

Four forms of analytics. Pita introduced the four forms of analytics (mentioned in Section 4.1), answering questions of *What happened?* (descriptive), *Why?* (diagnostic), *What will happen next?* (predictive), and *What should I do?* (prescriptive). Working with Armorway, Avata's customers are encouraged to incorporate as much data from the problem domain as possible and describe its ontology. The system can absorb data in a range of formats, apply transformations and cleanups, and take context into account.

Different kinds of data contribute to different analytics. For example, for descriptive analytics, Avata provides enhanced search methods in semi-structured and unstructured data. Information about a specific street address is obtained by combining data sources including maps, social media and review sites, web pages, maps, closed-circuit television footage, and so on.

For diagnostic analytics, a typical example is to look for correlations between features in different data sets. Pita made the point that having more data helps avoid identifying spurious correlations (correlation without causation). He gave an example of a spurious correlation between ice cream consumption and swimming pool drowning. Of course it is not ice cream that leads to the increase in drownings, but rather the hot temperatures that increase swimming pool use. Incorporating weather information into the set of data managed by the system allows the correct correlation to be inferred. Sometimes correlations are hard to discover and must be expressed to the system, using *objectives* to direct it. In law enforcement, an objective might be thefts. Then Armorway can search for possible *indicators* (e.g., vandalism) and refine these in a feedback loop. More generally, feedback can help steer the embedded ML algorithms ("thumbs up" and "thumbs down" for supervised training). Ultimately, incorporating histories of events, the system might be asked questions such as, *Does a full moon coincide with increased crime rates?*

Predictive and prescriptive analytics are more complex. For the prescriptive case, Avata notes that objectives may change over time. Using game-theory-based techniques allows their system to consider the best course of action.

Making predictions. For predictive analysis, *behavioral modeling* is needed. Choice theory is one technique used here; it provides a foundation to understand the motivation of people and it can work with different amounts of data. To deal with the problem of an infinite set of possible futures, they try to reduce to a set of *representative futures* the things that are most likely to happen next. Pita used the analogy of an "investment portfolio" to explain a set of possible actions that might lead to different results; then, the aim is to balance the portfolio to maximize the overall outcome. Rather than making "point predictions", the idea is to give a range of possible actions that should be robust.

Response. Their platform provides an alerting system and the ability to display information overlays on data. An example is icons added to a map of crime locations (shown in Fig. 7). The aim is to minimize the number of lower-severity alerts. At the time of the meeting, they were entering into discussions with security companies and considering further kinds of threat mitigation responses (e.g., taking systems offline). Pita presented some screens from the user interface of Armorway as well as its output on some projects.

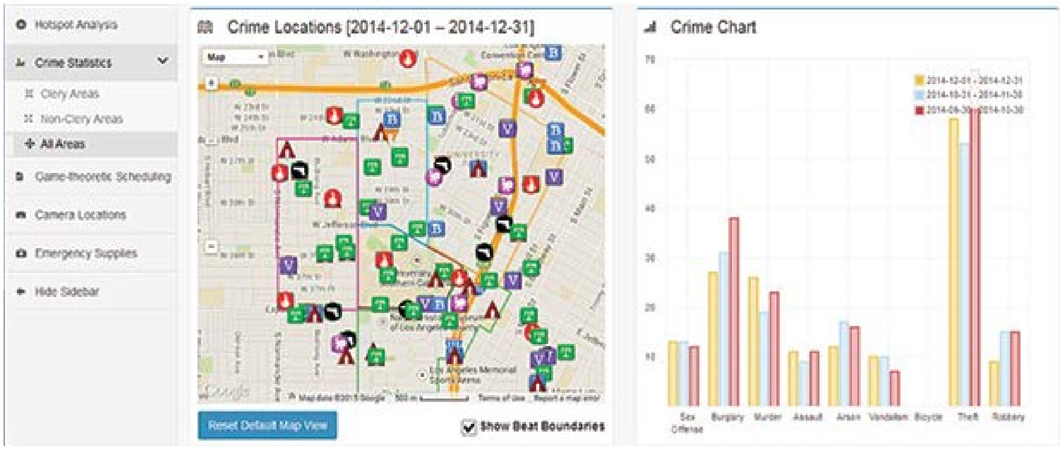


Fig. 7 Image from Security Magazine April 1, 2015 article, *Uses game theory to enhance security patrols and programs*

Discussion. Pita's presentation led to a lively discussion. Many questions were asked concerning how the results of such a system could be put to the test; how decisions might be made explainable to users (here, *indicators* might be used); whether adding continually increasing amounts of data would necessarily improve outcomes; how to choose the best methods for different problems; what the false positive and false negative rates are like; and whether so-called "black swan" events or low-rate attacks could also be modeled. The choice of methods appears to be not fully automated in the system but managed during employment through expert consultants.

Summary. Wide-spectrum AI toolkits like Avata's platform clearly have the opportunity to be transformative when configured to work on bespoke problem domains. In cybersecurity alone, many companies have been founded in the past five years selling tools deploying AI techniques to different ends and it has been a hot area for venture capital funding. However, the underlying problems of prediction in the cyber domain, although appealing to work on, can be especially demanding and likely requires more fundamental research. As Pita said, one of the key challenges is measurement: to see the effect of some mitigations one would like to be able to switch off (or compare with another data set where it is not used), but this generally cannot be done.

8.2 DHS Chief Information Officer (CIO) John Zangardi and Chief Information Security Officer (CISO) Paul Beckman

We invited presentations from the DHS's CIO, John Zangardi, and CISO, Paul Beckman, on the topic of predictive analysis of adversarial cybersecurity.

Beckman opened by saying that we know cybersecurity is the emerging battle of the 21st century and chief among his concerns is how to hire the experts needed to fight the cyber battle. Zangardi said that given the increasing importance of securing cyber and IT talent to his agency's mission, the federal hiring process desperately had to become more flexible to be competitive with the private sector. Paraphrasing, he said:

> "I am competing with the private sector on salary, and I'm competing with them on quality of life, while not having an easier hiring process. But the thing I am really competing on—and this is where I think I beat industry—is service to the nation and mission. I've heard that a lot from my buddies in industry, that they just don't get that job satisfaction. Job satisfaction is where I can win, but where I can't always win is on salary; the salary differential is really high."

We have heard similar sentiments from other national security agencies. The DHS is in the midst of a multipronged modernization strategy that involves projects like leveraging the Enterprise Infrastructure Solutions contract for network upgrades and telecommunication replacements, data center and security operations center consolidations, and other enhancements, all of which will require increasing numbers of IT professionals to drive progress.

Zangardi said because technology talent is critical to DHS's plans for both cybersecurity and IT modernization, his 2019 budget includes cyber pay compensation increases to help narrow the skills gap across the department and retain the talent it currently has to the private sector or other agencies. That will help him compete within the broader cyber community to retain staff; at the moment, retention is a problem because they may leave after training (even to other government departments).

DHS told us about international trips they made recently, including a delegation from DHS and various component agencies to meet with companies and government organizations from South Korea and Japan for ideas around various emerging technologies. The purpose of their visit was to learn about 5G, the next-generation wireless network, internet of things, cybersecurity, and ML. In Korea, they visited the US Embassy in South Korea and various Japanese government agencies, including the National Center of Incident Readiness and Strategy for

Cyber Security. DHS officials have also scheduled visits to Samsung, Hitachi, Docomo, Honda Robotics, and Nippon Telegraph and Telephone Data. The objective is to learn about how these companies and organizations are moving forward in cybersecurity.

In total, DHS spends more than $6 billion annually on IT. Traditionally, DHS has spent more than 85% of its IT budget on operating and maintaining traditional legacy systems, which is an increasing burden for security.

Tech envoys are not uncommon, especially as cabinet leaders and federal officials grapple with how to handle and implement rapidly evolving technologies. Last year, for example, Defense Secretary Jim Mattis and other Defense officials visited the headquarters of Google and Amazon—a trip that preceded a major Defense cloud procurement. The cyber threats are the number one focus for DHS. They are growing in complexity, volume, and frequency.

The DHS information security seniors agreed "we are not winning this battle just yet" and were interested to learn about NATO RTG activities on specific subproblems. Generally, with technology, products and cyber-space itself being so geographically distributed, international discussion is essential to make progress.

8.3 Amazon Web Services: Hayes Magnuson

Hayes Magnuson of AWS presented to us. Cloud security at AWS is the highest priority. Their aim is that AWS customers will benefit from a data center and network architecture built to meet the requirements of the most security-sensitive organizations.

An advantage of the AWS cloud is that it allows customers to scale and innovate, while maintaining a secure environment. Customers pay only for the services they use, meaning that a customer can have the appropriate security they need but without the upfront expenses, and at a lower cost than in an on-premises environment.

AWS recognizes the additional level of effort an organization has to expend for each new security assurance framework it implements. To reduce that burden, they provide a detailed breakout of AWS cloud offerings and associated customer and AWS responsibilities to facilitate alignment with the NIST Cybersecurity Framework (CSF) (Amazon 2019). Organizations ranging from federal and state agencies to regulated entities and large enterprises use the white paper as a guide for implementing AWS solutions to achieve the CSF risk management outcomes. AWS aims to provide data protection assurances and provide resources to secure their environments.

Although not discussed explicitly, we expect that in future, predictive analytics frameworks may be part of the offering from cloud service providers. It will obviously be important for many of their customers to understand whether they are likely to be at risk from attack in the future (and from which quarters), as well as understanding the fundamental limits of the security provision possible for a cloud provider to give. Even with the most effective and efficient police forces operating on the street, we still need to take care to lock our homes and automobiles.

8.4 Extreme Networks: John Szewc

Extreme Networks is a networking company based in San Jose, California, focusing on software-driven networking solutions for enterprise and service provider customers. They gave us a presentation discussing ML and AI in cybersecurity, which we present as a series of questions and answers from the company.

Question 1: *How do you define machine learning and artificial intelligence?*

ML is largely based on statistical mechanics and analysis. ML can take in a lot of big information and run mathematical algorithms and disseminate knowledge to provide a clear view of what is going on. Using machines allows us humans to make decisions around the data that is generated. Machines can process more data and process it faster than humans can, which is why ML is mainstream and widely used today.

AI, on the other hand, is at the cutting edge. Ultimately, AI requires a system that will start up and basically learn its environment the way a human does. In this view, AI is virtually nonexistent today; not even IBM Watson is artificially intelligent.

Question 2: *How important is the knowledge base that feeds AI/ML?*

Very important. Information is broad—some information can be represented well statistically and some information cannot. The construct of the knowledge base is a crucial factor. The narrower the construct of information, the more suited it is for ML and AI.

Consider a game of chess; it has a narrow, well-defined set of rules. IBM Watson, a question-answering computer system capable of answering questions posed in natural language, can play the world's top chess player and win. The reason is the knowledge base of information is well defined; it is easy to "train" Watson in every possible move. In contrast, using Watson to diagnose precancerous tumors has been proven to be less effective. Why? Because the knowledge base of information is less defined. Even the world's best research scientists and doctors do not have all the answers. In this environment, the scientists can more accurately diagnose tumors than Watson because in addition to relying on data, they leverage their experience and intuition. This does not mean that Watson is useless, it just means that it falls short of the definition of true AI in this space.

The key point is that machines need to be trained, like humans do, and lack the flexibility in ontology that humans have. Think about it. We can effectively move between a vast scope of environments with minimal training. Machines cannot do this as effectively. They need to be trained to a problem or action. Do not underestimate the power of human intuition.

Question 3: *Do you see a role for ML and AI in cybersecurity?*

Absolutely. There are many areas where ML solutions are being effectively used today. The security environment is constantly changing and evolving and ML solutions work best in areas where there is a narrow construct of information.

Threat detection systems and firewalls are a good example. Consider someone in Europe getting hacked by a new method: botnet or malware. That event is registered, a signature is uploaded to the cloud so that when someone in the US gets hacked, the breach has been seen before. In this instance, cloud technology and ML works like the human immune system—recognizing and reacting to the threat.

Another example is Active Directory (AD). By correlating AD logs, you can determine brute force enumeration attacks, impersonations, administrative account privilege searches, and so on. The challenge is the sheer volume of AD information makes it difficult to tease out what is important. ML can take a lot of garbage out of the way to allow humans to focus on the data sets that are important. We call this "actionable knowledge".

Question 4: *Any concerns about the negative use of AI/ML in cybersecurity?*

Yes! We are already seeing examples where criminals are taking advantage of AI/ML solutions. If I can manipulate the knowledge base on which security systems are making decisions on, I can compromise it. This is why it is essential never to rely solely on technology—you always need human involvement.

Consider self-driving cars and the danger of manipulating the knowledge base. What would happen if the traffic signal or stop sign suddenly disappeared from an autonomous car's view? It could have dire consequences. However, these types of things can, and do, happen in the cyber world. You need to be vigilant in protecting the integrity of your knowledge base and you can never take the human completely out of the loop. Similar to how a human can regain control of the car, there will be times where you need a human to regain control of security.

Question 5: *Where should enterprises start when it comes to AI/ML?*

My advice is twofold. First, invest in information: 90% of what we talked about is information. The more information that you have that can be correlated against one another the better. Invest in systems that can work in an open ecosystem. Look for solutions with end-to-end analytics capabilities, application telemetry, and so on.

Second, invest in security experts. We will never take humans out of the mix. You cannot rely on ML/AL alone, so invest in human knowledge and give them the right tools. Look to hire what you do not have: penetration experts, risk assessment experts, and so on.

Summary. Overall, cybersecurity was a top priority for Extreme Networks and they participated in initiatives such as the Openflow Consortium to support standard mechanisms for reporting as well as control. As IT departments are moving more and more of their business into the cloud, they expect to work with cloud providers to help develop their cloud networking cyber strategies.

9. Current State of the Art and Survey

One of the keys items the RTG addressed was the need for a current survey of the state of the art with respect to prediction of adversarial behavior. This section of the report is meant to provide some insight into the issues the committee faced in trying to appropriately characterize our research and to also provide some data on our experience with what actually existed in the general body of knowledge.

The committee spent the first seven meetings discussing, rehashing, editing, and using sets of "key words" to provide us with foundational and related research. Our initial expectation was that there should be collateral papers and related research that could possibly be beneficial to helping shape our research. We expected to be able to search for prediction in general and then use that as a guide to refine the results. As the committee found out, there was a lack of research we considered foundational.

9.1 Formulating Literature Search Parameters

This section provides some insight into the process of selecting and then refining search parameters that the RTG used. We captured that here because we believe it will be useful for future research in this area.

We had a crowdsourcing exercise and developed a list of what we considered reasonable criteria that would describe characteristics of a solution to the cyber predictive problem. The list that follows represents the list of search terms based on facets of a solution:

- Fusion of different analytic approaches (implementation, classes): This represents a realization that a potential solution would require an integrated set of analytics.
- Situation description methods, for example, elements, features, and interrelationship between components: What prior research was done that had bearing on using different analytics to solve similar problems based on the prediction environment prediction assumptions and available parameters?
- Prediction methods: What research has been done on methods of prediction and how successful were they?
- Detection as a function of threat capability (as discussed earlier from the DSB report [Gosler and Von Thaer 2013]): Has there been research done based on threat capabilities as illustrated in Table 1?
- Cyber-attack probability (indications and warnings [I&W]): The RTG was aware of some early research work on cyber I&W. We felt a review of the current state of that practice was in order.
- Evaluation of threats that leverage known vulnerabilities with previously unseen exploits. Related to the DSB report, but specifically looking for new exploits attacking known vulnerabilities.
- Attack graph generation of known vulnerabilities/unknown exploits: Related to the previous bullet, has there been research or success using attack graphs to identify new exploits attacking known vulnerabilities?
- The combination attack graphs/models/paths *and* intrusion detection: The RTG thought the intersection of attack graph theory and methods of intrusion detection could yield some knowledge on real-time usage of attack graphs for intrusion detection.
- Cyber attacker profiling/TTPs: Use of reports similar to Krekel (2009) and Krekel et al. (2014).
- M&S for generation of adversarial TTPs: The RTG was aware of areas within NATO that were beginning to look at modeling in depth for the cyber environment. This search would evaluate how far that has advanced in the body of knowledge.

- Identification of adversarial behavior: What parameters constitute adversarial behavior and have any been defined for the cyber domain?
- COAs as a function of threat capability: Have systems been deployed or testing done to create cyber COAs against an imminent or ongoing cyber-attack?
- Detection of unknown vulnerabilities: What automated or semi-automated algorithmic approaches have been used to identify unknown vulnerabilities within a system? Similarly, what process do "cyber warriors" use to identify unknown vulnerabilities?
- Influence and knowledge on risk management of cyber COAs.
- Reevaluating the results of the crowdsourced exercise, the committee re-formed the list:
    - Adversary behavior identification—captured under attacker characterization
    - Attack graph generation of known vulnerabilities/unknown exploits
    - Attack probability I&Ws—Under prediction for nonsignature-based I&Ws
    - Combination detection analytics
    - Cyber OODA loop
    - Attacker characterization vs. TTPs
    - Fusion of different analytic approaches for prediction of nonsignature-based cyber attacks
    - M&S of cyber COAs
    - Papers on multistep attack models (~attack graphs) and intrusion detection
    - Prediction methods in other domains—The committee considered this a subset of prediction technology and implementation, so it was removed
    - Prediction technology and implementation
    - Reference documents
    - Risk management and cyber COAs

    - Situation description methods and tools
    - Threat COAs vs. capabilities
    - Threat detection and COAs vs. capability
    - Threats using known vulnerability-unknown exploits

After performing some literature searches during Meetings 4 and 5 and finding not as much material as we hoped, we streamlined and combined the parameters:

- Cyber profiling vs. TTPs that included threat characterization, attacker characterization, and adversary behavior identification
- Fusion of different analytic approaches for prediction of nonsignature-based cyber-attacks
- Prediction technology and implementation that also included attack probability indications and warnings, and prediction methods in other domains
- Risk management cyber COAs
- Situation description methods and tools
- Threat detection and COAs vs. capability that also included M&S, threats using known vulnerability-unknown exploits, attack graph generation of known vulnerabilities/unknown exploits, and papers on multistep attack models (~attack graphs) and intrusion detection

A partial list of papers we evaluated through the three exercises described previously is contained in Table 7. We strongly believe that a great deal of work has been done attempting to examine the problem of threat prediction. We were essentially searching for approaches that predicted next behavior given few, if any, constraints. We were not looking for prediction between established outcomes, rather we looked for prediction of what those outcomes might be. The list of 100 references illustrates the examination at one of our 10 meetings. All of these papers, while not specifically addressing our definition of prediction, are important papers in the development of cohesive cyber defense.

Table 7 Partial list of cyber defense papers evaluated by IST-129

Paper no.	Reviewed cyber defense papers
1	Albertson D. Visual information seeking. Journal of the Association for Information Science and Technology, 2015;66(6):1091–1105. https://doi.org/10.1002/asi.23244.
2	Andress J, Winterfeld S. Cyber warfare: techniques, tactics and tools for security practitioners. Elsevier; 2013.
3	Applegate SD. The principle of maneuver in cyber operations. In: 2012 4th International Conference on Cyber Conflict (CYCON 2012) (p. 1-13). IEEE; 2012 June.
4	Azuma R, Daily M, Furmanski C. A review of time critical decision making models and human cognitive processes. In: 2006 IEEE aerospace conference (p. 9). IEEE; 2006 Mar.
5	Bean J. Characterization of relevant attributes using cyber trajectory similarities. 2009.
6	Bell B, Santos Jr E, Brown SM. Making adversary decision modeling tractable with intent inference and information fusion. In: Proceedings of the 11th Conference on Computer Generated Forces and Behavioral Representation; 2002 May; Orlando, FL.
7	Ben-david R. Enhancing comprehension through graphic organizers. 2002.
8	Berral JL, Poggi N, Alonso J, Gavalda R, Torres J, Parashar, M. Adaptive distributed mechanism against flooding network attacks based on machine learning. In: Proceedings of the 1st ACM workshop on Workshop on AISec (p. 43–50); 2008 Oct.
9	Bilge L, Dumitraş T. Before we knew it: an empirical study of zero-day attacks in the real world. In: Proceedings of the 2012 ACM Conference on Computer and Communications Security (p. 833–844); 2012 Oct.
10	Boddy MS, Gohde J, Haigh T, Harp SA. Course of action generation for cyber security using classical planning. In: ICAPS (p. 12–21); 2005 June.
11	Bozorgi M, Saul LK, Savage S, Voelker GM. Beyond heuristics: learning to classify vulnerabilities and predict exploits. In: Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (p. 105–114); 2010 July.
12	Brehmer B. The dynamic OODA loop: amalgamating Boyd's OODA loop and the cybernetic approach to command and control. In: Proceedings of the 10th International Command and Control Research Technology Symposium (p. 365–368); 2005 June.
13	Bryant DJ. Rethinking OODA: toward a modern cognitive framework of command decision making. Military Psychology. 2006;18(3):183–206.
14	Byers SR, Yang SJ. Real-time fusion and projection of network intrusion activity. In: 2008 11th International Conference on Information Fusion (p. 1–8). IEEE; 2008 June.
15	Canali D, Bilge L, Balzarotti D. On the effectiveness of risk prediction based on users browsing behavior. In: Proceedings of the 9th ACM Symposium on Information, Computer and Communications Security (p. 171–182); 2014 June.
16	Carter KM, Idika N, Streilein WW. Probabilistic threat propagation for network security. IEEE Transactions on Information Forensics and Security. 2014;9(9):1394–1405.

Table 7 Partial list of papers evaluated by IST-129 (continued)

Paper no.	APA citation reference
17	Champion MA, Rajivan P, Cooke NJ, Jariwala S. Team-based cyber defense analysis. 2012 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support, CogSIMA 2012; 2012 March, p. 218–221. https://doi.org/10.1109/CogSIMA.2012.6188386.
18	Chen HM, Kazman R, Monarch I, Wang P. Predicting and fixing vulnerabilities before they occur: a big data approach. In: Proceedings of the 2nd International Workshop on BIG Data Software Engineering; 2016 May; (p. 72–75).
19	Cheng BC, Liao GT, Huang CC, Yu T. A novel probabilistic matching algorithm for multi-stage attack forecasts. IEEE Journal on Selected Areas in Communications. 2011;29(7):1438–1448.
20	Conti G, Nelson J, Raymond D. Towards a cyber common operating picture. In: 2013 5th International Conference on Cyber Conflict (CyCon) (p. 1–17). IEEE; 2013 June.
21	Colbaugh R, Glass K. Predictive defense against evolving adversaries. In: 2012 IEEE International Conference on Intelligence and Security Informatics (p. 18–23). IEEE; 2012 June.
22	Cybenko G. Cyber adversary dynamics. Dartmouth Coll Hanover NH Thayer School of Engineering; 2013.
23	D’Amico AD, Goodall JR, Tesone DR, Kopylec JK. Visual discovery in computer network defense. IEEE Computer Graphics and Applications. 2007;27(5):20–27. https://doi.org/10.1109/MCG.2007.137.
24	D’Amico A, Kocka M. Information assurance visualizations for specific stages of situational awareness and intended uses: lessons learned. IEEE Workshop on Visualization for Computer Security 2005, VizSEC 05, Proceedings, p. 107–112; 2005. https://doi.org/10.1109/VIZSEC.2005.1532072.
25	Das S, Mukhopadhyay A, Shukla GK. i-HOPE framework for predicting cyber breaches: a logit approach. In 2013 46th Hawaii International Conference on System Sciences (p. 3008–3017). IEEE; 2013 Jan.
26	Degeler V, French R, Jones K. Self-healing intrusion detection system concept. In: 2016 IEEE 2nd International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS) (p. 351–356). IEEE; 2016 Apr.
27	Dillon T, Chang E. Trust and risk semantics and prediction through big data analytics to encompass cloud services, cyber-physical systems, and social media: issues and challenges. In: 2016 IEEE Trustcom/BigDataSE/ISPA (p. 185–193). IEEE; 2016 Aug.
28	Dong X, Li Y, Wei S. Design and implementation of a cognitive engine functional architecture. Chinese Science Bulletin. 2012;57(28-29):3698–3704.
29	Eom JH, Kim NU, Kim SH, Chung TM. Cyber military strategy for cyberspace superiority in cyber warfare. In: Proceedings 2012 International Conference on Cyber Security, Cyber Warfare and Digital Forensic, CyberSec 2012 (p. 295–299); 2012. https://doi.org/10.1109/CyberSec.2012.6246114.
30	Farhadi H, AmirHaeri M, Khansari M. Alert correlation and prediction using data mining and HMM. ISeCure-The ISC International Journal of Information Security. 2011;3(2):77–101.

Table 7 Partial list of papers evaluated by IST-129 (continued)

Paper no.	APA citation reference
31	Fava D, Holsopple J, Yang SJ, Argauer B. Terrain and behavior modeling for projecting multistage cyber attacks. In: 2007 10th International Conference on Information Fusion (p. 1–7). IEEE; 2007 July.
32	Fava DS, Byers SR, Yang SJ. Projecting cyberattacks through variable-length markov models. IEEE Transactions on Information Forensics and Security. 2008;3(3):359–369.
33	Feng J, Yuan Z, Yao S, Xia C, Wei Q. Generating attack scenarios for attack intention recognition. In 2011 International Conference on Computational and Information Sciences (p. 272–275). IEEE; 2011 Oct.
34	Franke U, Brynielsson J. Cyber situational awareness – a systematic review of the literature. Computers & Security. 2014;46:18–31. https://doi.org/10.1016/j.cose.2014.06.008.
35	Frazier P, Lin R, McCallam D. Examining correlation techniques to improve strategic decision-making through advanced cyber situational awareness.
36	Geers K, Kindlund D, Moran N, Rachwald R. World War C: understanding nation-state motives behind today's advanced cyber attacks. FireEye, Milpitas, CA, USA, Tech. Rep., 2014 Sep.
37	Gilmore DA, Krause LS, Lehman LA, Santos Jr E, Zhao Q. Intent driven adversarial modeling. Air Force Research Lab, Rome NY; 2005.
38	Gray D. Improving cybersecurity governance through data-driven decision-making and execution (briefing charts). Carnegie-Mellon Univ Pittsburgh PA Software Engineering Inst; 2014.
39	Greitzer FL, Frincke DA. Combining traditional cyber security audit data with psychosocial data: towards predictive modeling for insider threat mitigation. In: Insider Threats in Cyber Security (p. 85–113). Springer; Boston, MA; 2010.
40	Harman D, Brown S, Henz B, Marvel LM. (2015). A communication protocol for CyAMS and the cyber fighter associate interface. Army Research Laboratory (US); 2015. Report No.: ARL-TN-0673.
41	Hamilton SN. Automated adversary profiling. In: Cyber Warfare (p. 141–149). Springer, Cham; 2015.
42	Heckman KE, Stech F. Cyber counterdeception: how to detect denial & deception (D&D). In: Cyber Warfare (p. 103–140). Springer, Cham; 2015.
43	Holm H. A framework and calculation engine for modeling and predicting the cyber security of enterprise architectures. (Doctoral dissertation, KTH Royal Institute of Technology); 2014.
44	Hutchins EM, Cloppert MJ, Amin RM. Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. Leading Issues in Information Warfare & Security Research. 2011;1(1):80.
45	Jaganathan V, Cherurveettil P, Muthu Sivashanmugam P. Using a prediction model to manage cyber security threats. The Scientific World Journal. 2015.
46	Kanoun W, Cuppens-Boulahia N, Cuppens F, Dubus S, Martin, A. Success likelihood of ongoing attacks for intrusion detection and response systems. In: 2009 International Conference on Computational Science and Engineering (Vol. 3, p. 83–91). IEEE; 2009 Aug.
47	Karaman M, Catalkaya H, Gerehan AZ, Goztepe K. Cyber operation planning and operational design. Cyber-Security and Digital Forensics. 2016;21.

Table 7 Partial list of papers evaluated by IST-129 (continued)

Paper no.	APA citation reference
48	Kim G, Lee S, Kim S. A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. Expert Systems with Applications. 2014;41(4):1690–1700.
49	Kime BP. Threat intelligence: Planning and direction. The SANS Institute [accessed March 17, 2017]; 2016.
50	Kott A, Ownby M. Toward a research agenda in adversarial reasoning: Computational approaches to anticipating the opponent's intent and actions. arXiv preprint arXiv:1512.07943;2015.
51	Kordy B, Piètre-Cambacédès L, Schweitzer P. DAG-based attack and defense modeling: Don't miss the forest for the attack trees. Computer Science Review. 2014;13:1–38.
52	Kotenko I, Chechulin A. A cyber attack modeling and impact assessment framework. In: 2013 5th International Conference on Cyber Conflict (CYCON 2013)(p. 1–24). IEEE; 2013 June.
53	Leed M. Offensive cyber capabilities at the operational level. Center for Strategic International Studies. Georgia Tech Research Institute. Washington, DC; 2013.
54	Lehto M, Neittaanmäki P. (eds.). Cyber security: analytics, technology and automation (Vol. 78). Springer; 2015.
55	Lei J, Li ZT. Using network attack graph to predict the future attacks. In: 2007 Second International Conference on Communications and Networking in China (p. 403–407). IEEE; 2007 Aug.
56	LeMay E, Ford M, Keefe K, Sanders WH, Muehrcke C. Model-based security metrics using adversary view security evaluation (advise). In: 2011 Eighth International Conference on Quantitative Evaluation of SysTems (p. 191–200). IEEE; 2011 Sep.
57	Lenders V, Tanner A, Blarer A. Gaining an edge in cyberspace with advanced situational awareness. IEEE Security & Privacy. 2015;13(2):65–74.
58	Lévesque FL, Fernandez JM, Somayaji A. Risk prediction of malware victimization based on user behavior. In: 2014 9th international conference on malicious and unwanted software: The Americas (MALWARE) (pp. 128-134). IEEE; 2014, October.
59	Li W, Zhi-tang L, Qi-hong W. A novel technique of recognizing multi-stage attack behaviour. In: 2006 International Workshop on Networking, Architecture, and Storages (IWNAS '06) (p. 188–193). IEEE; 2006 Aug.
60	Liu Y, Sarabi A, Zhang J, Naghizadeh P, Karir M, Bailey M, Liu, M. Cloudy with a chance of breach: forecasting cyber security incidents. In: 24th {USENIX} Security Symposium ({USENIX} Security 15) (p. 1009–1024);2015.
61	Liu Y, Zhang J, Sarabi A, Liu M, Karir M, Bailey M. Predicting cyber security incidents using feature-based characterization of network-level malicious activities. In: Proceedings of the 2015 ACM International Workshop on International Workshop on Security and Privacy Analytics (p. 3–9); 2015 Mar.
62	Magoutas B, Stojanovic N, Bousdekis A, Apostolou D, Mentzas G, Stojanovic L. Anticipation-driven architecture for proactive enterprise decision making. In: CAiSE (Forum/Doctoral Consortium) (p. 121–128);2014.

Table 7 Partial list of papers evaluated by IST-129 (continued)

Paper no.	APA citation reference
63	Mahmood T, Afzal U. Security analytics: big data analytics for cybersecurity: a review of trends, techniques and tools. In: 2013 2nd National Conference on Information Assurance (NCIA) (p. 129–134). IEEE; 2013 Dec.
64	McCallam D. An analysis of cyber reference architectures. Presented at NATO 2012 Workshop with Industry on Cybersecurity Capabilities; 2012.
65	McCallam DH, Frazier PD, Savold R. Ubiquitous connectivity and threats: architecting the next generation cyber security operations. In: 2017 IEEE 7th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER) (p. 1506–1509). IEEE; 2017 July.
66	Medvedev SA. Offense-defense theory analysis of Russian cyber capability. Naval Postgraduate School; Monterey, CA;2015.
67	Mitchell W. Battlespace agility 201: the OODA moment; 2013.
68	Mushtaq MT, Khan MS, Naqvi MR, Khan RD, Khan MA, Koudelka OF. Cognitive radios and cognitive networks: a short introduction. Journal of Basic & Applied Scientific Research; 2013.
69	Newmeyer KP. Elements of national cybersecurity strategy for developing nations. National Cybersecurity Institute Journal. 2015;1(3):9–19.
70	Noel S, Robertson E, Jajodia S. Correlating intrusion events and building attack scenarios through attack graph distances. In :20th Annual Computer Security Applications Conference (p. 350–359). IEEE; 2004 Dec.
71	Noel S, Jajodia S. Optimal IDS sensor placement and alert prioritization using attack graphs. Journal of Network and Systems Management. 2008;16(3):259–275.
72	Noel S, Ludwig J, Jain P, Johnson D, Thomas RK, McFarland J, Tello B. Analyzing mission impacts of cyber actions (AMICA). In: NATO IST-128 Workshop on Cyber Attack Detection, Forensics and Attribution for Assessment of Mission Impact; 2015.
73	Omarova A, Ireland V, Gorod A. An alternative approach to identifying and appraising adaptive loops in complex organizations. Procedia Computer Science. 2012;12:56–62.
74	Ou X, Rajagopalan SR, Sakthivelmurugan S. An empirical approach to modeling uncertainty in intrusion analysis. In: 2009 Annual Computer Security Applications Conference (p. 494–503). IEEE; 2009 Dec.
75	Qin X, Lee W. Attack plan recognition and prediction using causal networks. In: 20th Annual Computer Security Applications Conference (p. 370–379). IEEE; 2004 Dec.
76	Raska M. Decoding China's cyber warfare strategies. 2015.
77	Rausch M, Feddersen B, Keefe K, Sanders WH. A comparison of different intrusion detection approaches in an advanced metering infrastructure network using ADVISE. In: International Conference on Quantitative Evaluation of Systems (p. 279–294). Springer, Cham; 2016 Aug.
78	Rieck K, Laskov P. Language models for detection of unknown attacks in network traffic. Journal in Computer Virology. 2007;2(4):243–256.
79	Roschke S, Cheng F, Schuppenies R, Meinel C. Towards unifying vulnerability information for attack graph construction. In: International Conference on Information Security (p. 218–233). Springer, Berlin, Heidelberg; 2009 Sep.

Table 7 Partial list of papers evaluated by IST-129 (continued)

Paper no.	APA citation reference
80	Santos Jr E. A cognitive architecture for adversary intent inferencing: Structure of knowledge and computation. In: Enabling Technologies for Simulation Science VII (Vol. 5091, p. 182–193). International Society for Optics and Photonics; 2003 Sep.
81	Santos Jr E, Zhao Q. Adversarial models for opponent intent inferencing. In: Kott AW, McEneaney WM, editors. Adversarial reasoning: computational approaches to reading the opponents mind. CRC Press; c2006. p. 1–22.
82	Savold R, Dagher N, Frazier P, McCallam D. Architecting cyber defense: a survey of the leading cyber reference architectures and frameworks. In: 2017 IEEE 4th International Conference on Cyber Security and Cloud Computing (CSCloud) (p. 127–138). IEEE; 2017 June.
83	Sheyner O, Wing J. Tools for generating and analyzing attack graphs. In: International Symposium on Formal Methods for Components and Objects (p. 344–371). Springer, Berlin, Heidelberg; 2003 Nov.
84	Snyder D, Hart GE, Lynch KF, Drew JG. Ensuring US Air Force operations during cyber attacks against combat support systems: guidance for where to focus mitigation efforts. Rand Project Air Force; Santa Monica, CA; 2015.
85	Sommestad T, Sandström F. An empirical test of the accuracy of an attack graph analysis tool. Information & Computer Security. 2015.
86	Stech FJ, Heckman KE, Strom BE. Integrating cyber-D&D into adversary modeling for active cyber defense. In: Cyber Deception (p. 1–22). Springer, Cham; 2016.
87	Suthaharan S. Big data classification: problems and challenges in network intrusion prediction with machine learning. ACM SIGMETRICS Performance Evaluation Review. 2014;41(4):70–73.
88	Swanson S, Astrich C, Robinson M. Cyber threat indications & warning: predict, identify and counter. Journal Article. July 2012;26(4):59am.
89	Vamvoudakis KG, Hespanha JP, Kemmerer RA, Vigna G. Formulating cyber-security as convex optimization problems. In: Control of Cyber-Physical Systems (p. 85–100). Springer, Heidelberg; 2013.
90	Veerasamy N. High-level mapping of cyberterrorism to the OODA loop. 2010.
91	Veeramachaneni K, Arnaldo I, Korrapati V, Bassias C, Li K. AI^ 2: training a big data machine to defend. In: 2016 IEEE 2nd International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS) (p. 49–54). IEEE; 2016 Apr.
92	Wang L, Jajodia S, Singhal A, Cheng P, Noel S. k-zero day safety: a network security metric for measuring the risk of unknown vulnerabilities. IEEE Transactions on Dependable and Secure Computing. 2013;11(1):30–44.
93	Wechsler H. Cyberspace security using adversarial learning and conformal prediction. Intelligent Information Management. 2015;7(04):195.
94	Williamson ML. The cyber military revolution and the need for a new framework of war. National Defense Univ Norfolk VA Joint Advanced Warfighting School; 2012.

Table 7 Partial list of papers evaluated by IST-129 (continued)

Paper no.	APA citation reference
95	Wu J, Yin L, Guo Y. Cyber attacks prediction model based on Bayesian network. In: 2012 IEEE 18th International Conference on Parallel and Distributed Systems (p. 730–731). IEEE; 2012 Dec.
96	Yang SJ, Stotz A, Holsopple J, Sudit M, Kuhl M. High level information fusion for tracking and projection of multistage cyber attacks. Information Fusion. 2009;10(1):107–121.
97	Yuen J, Turnbull B, Hernandez J. Visual analytics for cyber red teaming. 2015 IEEE Symposium on Visualization for Cyber Security; VizSec 2015. 2015 Nov. doi.org/10.1109/VIZSEC.2015.7312765.
98	Zhang S, Zhang X, Ou X. After we knew it: empirical study and modeling of cost-effectiveness of exploiting prevalent known vulnerabilities across iaas cloud. In: Proceedings of the 9th ACM Symposium on Information, Computer and Communications Security (p. 317–328); 2014 June.
99	Zhong C, Samuel D, Yen J, Liu P, Erbacher R, Hutchinson S, Glodek W. Rankaoh: context-driven similarity-based retrieval of experiences in cyber analysis. In: 2014 IEEE International Inter-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA) (p. 230–236). IEEE; 2014 Mar.
100	Zhu B, Ghorbani AA. Alert correlation for extracting attack strategies. IJ Network Security. 2006;3(3):244–258.

9.2 Discussions of Binning/Grouping Similar Papers

Paper validity was evaluated by RTG and results are shown in Fig. 8. All papers are newer than year 2000. The majority of valid papers are between 2006–2010.

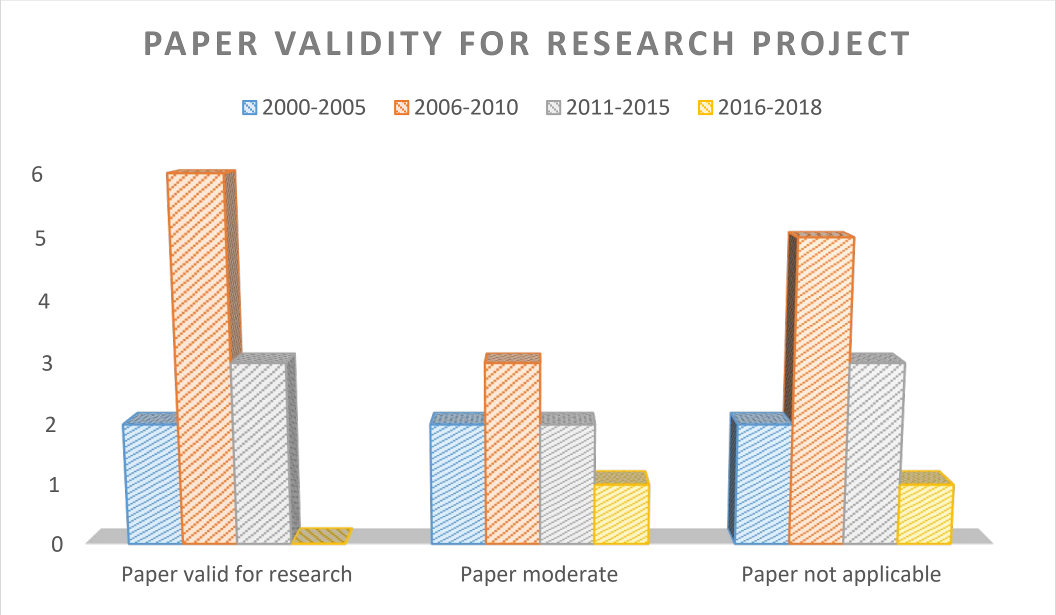


Fig. 8 How we performed the literature search

9.3 Information Extraction

The literature addressed the prediction problem from a number of different angles and at different levels of abstraction. At first, an input–output perspective attempted to broadly characterize the data the models used. The model of STIX (Barnum 2012) was used for this purpose. However, the analysis soon showed that most of the models used more or less the same data objects in STIX—namely, vulnerability information, attack patterns, indicators, intrusion sets, and other observed data. In addition, it turned out to be nontrivial to classify the data used in a reliable manner, partly because of the different levels of abstraction used in the papers. Instead, the following information was extracted to characterize the models:

1) If a particular formalism was used or proposed.

2) Data used as input for the prediction model.

3) The data produced as output by the prediction model.

4) The scalability of the solution or implementation.

To additionally characterize the models, it was extracted how they handled the following issues:

1) Adversaries attempting to tamper with/fool the prediction method.

2) The time it takes to make the prediction and timing issues.

3) Availability of data needed for analysis or model construction.

4) Assumptions concerning knowledge of system vulnerabilities and attacks.

Furthermore, information was extracted to characterize the maturity of the research in terms by assessing:

1) If the model had been implemented in prototype and what technology readiness level the model was on.

2) If the model's usefulness was demonstrated (e.g., in a case study).

3) Tests or other evaluations of accuracy of the prediction model.

These 11 information elements were extracted as quotes and summaries of descriptions provided in the final selection of all the reviewed papers:

1) Cheng BC, Liao GT, Huang CC, Yu MT. A novel probabilistic matching algorithm for multi-stage attack forecasts. IEEE J Sel Areas Commun. 2011;29:1438–1448. The authors propose a solution that inspects attack

graphs and guesses what the attackers are up to from the number of matched steps in the graph. Their method named JEAN (Judge Evaluation Attack iNtension) predicts possible attacks. The method is a probability-based approach. The authors demonstrate the method is more accurate and less labor intensive than LCS-(Longest Common Subsequence)-based approaches. However, they found out that is very hard to forecast multistage attacks.

2) Colbaugh R, Glass K. Predictive defense against evolving adversaries. ISI 2012. 2012 IEEE Int Conf Intell Secur Informatics Cyberspace, Border, Immigr Secur. 2012;18–23. This paper addresses prediction with ML and game theory. They apply it to spam to predict how spam evolves over time and tested it with real data. They use definitions of topological properties (transitivity, community structure, and core-periphery structure), which may have some direct translation to contextualizing cyber events.

3) Greitzer FL, Frincke DA. Combining traditional cybersecurity audit data with psychosocial data: towards predictive modeling for insider threat mitigation. In: Advances in Information Security. 2010;85–113. This paper focuses on insider threats and is about anticipating attacks from them. The predictions are about who will be malicious rather than when they will attack or how likely it is that someone will attack. The key to prediction is to incorporate traditional cyber-audit data with demographical and organizational data of the employee. They suggested that any data monitoring needed to predict should be based on actual behavior and events. They combine data fusion and analysis like predictive classification. To provide warning signs of cyber-attacks they suggested evaluating demographic, behavioral, and psychosocial data indicators based on case studies. The prediction should be tested against of a set of real cases. They gave usable, predictive indicators for developing a framework.

4) Lee S, Lee DH, Kim KJ. A conceptual design of knowledge-based real-time cyber-threat early warning system. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics) 4331 LNCS. 2006;1006–1017. This paper reviews several previous studies that tried to "predict" attacks. But "predicting" attacks here is just detecting unusual increases in the volume of network traffic using various techniques. One of the previous studies was a Kalman Filtering Forecast Model. The paper talks about an "early warning system" to warn administrators of a network attack. This is a form of prediction, but not really what we were looking for. Again, it's based mostly on the volume of network traffic going above an acceptable threshold. And the paper is 13 years old, so the makeup of the

network traffic is quite dated (they talk about detecting an MSN messenger attack).

5) Lei J, Li ZT. Using network attack graph to predict the future attacks. Proc Second Int Conf Commun Netw China. ChinaCom 2007. 2008;403–407. This paper is about making predictions. The authors use an attack graph with probability values tied to it but do not say how to obtain the probabilities. The authors create attack graphs from IDS data and predict attacks based on IDS events. Their experimental validation uses honeypots, which means it may not be representative.

6) Qin X, Lee W. Attack plan recognition and prediction using causal networks. Proc Annu Comput Secur Appl Conf ACSAC. 2004;370–379. The paper developed a graph-based technique to correlate isolated attack scenarios from isolated alerts. They converted attack trees to Bayesian networks and conducted probabilistic inference to evaluate the likelihood of attack goals and predict potential upcoming attacks. Some assumptions are made, and expert knowledge is required, but likely attacks on assets are predicted.

7) Santos E Jr. A cognitive architecture for adversary intent inferencing: knowledge structure and computation. Proc SPIE 17th Annu Int Symp Aerospace/Defense Sens Control. 2003;5091:182–193. This is an abstract paper about predicting adversaries' intent based on probability networks. For example, "case-based recognition" is discussed but the mathematical model is not presented in detail. The goal of the paper is to enable mission planning by using prediction for COAs.

8) Sarabi A, Bailey M. Predicting cybersecurity incidents using feature-based characterization of network-level malicious activities categories and subject descriptors. Int Work Secur Priv Anal (SPA '15). 2015;3–9. The article tries to predict cybersecurity incidents based on the assumption that IP address space prefixes have a small entropy of "badness". As a case study, they define a classifier based on blacklisted IP addresses and use this to predict a cyber-attack on a vulnerable network. Unfortunately, it is not clear if this is a valid assumption currently.

9) Shen D, Chen G, Blasch E, Tadda G. Adaptive Markov game theoretic approach for cyber network defense. MILCOM. 2007 Oct 29–31. The authors generate a primitive prediction of a cyber attacker's intents. High-level data fusion based on a Markov game model is proposed to capture new or unknown threats. The method is used to estimate the possible cyber-attack with uncertainty. Each player starts with some initial beliefs and

chooses the best response to those beliefs. They build a game-simulation platform and test the unknown threats through visualization and experiment.

10) Yang SJ, Byers S, Holsopple J, Argauer B, Fava D. Intrusion activity projection for cyber situational awareness. IEEE Int Conf Intell Secur Informatics. IEEE ISI 2008. 2008;167–172. Behavior trends for projections of future intrusions are based on Variable Length Markov Models; they put four elements for projecting cyber-attack actions (capability, opportunity, intent, and behavior). A conservative way is to assume that all attackers are able to execute all known and unknown exploitation methods and attackers used all services that they used before. Capability alone is not enough to estimate future attack actions. They developed a prototype of a virtual terrain model with algorithms for cyber-intrusion projection and they discovered that capability and opportunity are effective to project most cyber-attack actions. They used a 13-step attack and noted problems with decoy and stealthy attacks.

11) Yang SJ, Stotz A, Holsopple J, Sudit M, Kuhl M. High level information fusion for tracking and projection of multistage cyber attacks. Inf Fusion. 2009;10:107–121. This paper introduces information fusion to provide situation awareness and threat prediction. A fusion system is proposed for the tracking and projection of multistage attacks. The paper separates modeling of cyber-attack method from modeling of the network configuration. Predictions are performed independently on the two models, then fused to determine the targeted entities. They reference two previously existing tools/systems. First, they use INformation Fusion Engine for Real-time Decision-making (INFERD) to detect, correlate, and associate alerts that are part of multistage attack tracks. Second, Threat Assessment for Network Data and Information (TANDI) considers the current network status and results from these tools are combined to determine next most-likely targets. The system has several limitations, but this is one of the few papers we reviewed that provided actionable predictions.

10. NATO IST-145 Specialist Meeting

This section details the discussions and findings of the 2017 NATO IST-145 Specialist Meeting on Predictive Analytics and Analysis in the Cyber Domain held in Sibiu, Romania, 10–11 October 2017 at the Land Forces Academy. This Specialist Meeting was unclassified and open to NATO nations, Partner for Peace nations, Mediterranean Dialogue, Istanbul Cooperation Initiative nations, and Global Partners. The IST-145 Specialist Meeting is a derivative activity from the IST-129 NATO RTG on Predictive Analysis of Adversarial Cyber Operations. This section presents our findings prior to the Specialist Meeting, the background of the work from the RTG, a brief overview of the keynote address from Dr Eugene Santos from Dartmouth University (an expert in the field of prediction), and a summary of the Specialist Meeting's outcomes.

The proceedings of that workshop were published as a US Army Research Laboratory technical report (McCallam et al. 2019).

10.1 Findings Prior to the Specialist Meeting

Leading into the Specialist Meeting, we made some interim findings from our work:

1) The known vulnerability/known exploit (from the DSB reference on threat capability) is a solvable problem and has been solved, but not necessarily implemented through automation. It is detection as opposed to prediction, making prediction trivial in this case.

2) Prediction at the edge cases are outside the scope of effective prediction at this time. The edge values on the known vulnerability–unknown exploit capability threat (0% chance a cyber event will not occur and 100% certainty that a cyber event will occur) are potentially unattainable.

   a) Incidents can be independent variables and have no relation to previous cyber events. There is no guarantee that the sequence of cyber events identified represents a fully understood and known threat TTP.

   b) A prime example of this in real life are lottery games that present the occurrences of numbers in the previous draws tricking people into thinking the next draw is a function of previous draw(s).

   c) The IST Task Group felt that the Colin Powell credited quote, "As an intelligence officer, your responsibility is to tell me what you know. Tell me what you don't know. Then you're allowed to tell me what you think. But you always keep those three separated", has importance in the prediction process since this distinguishes between 100% prediction and likelihood. Not being this specific could have adverse effects on cyber-defensive positions.

3) The IST-129 Task Group felt a common taxonomy was needed to communicate in the cyber-prediction domain and recommends the use of STIX as a consistent means of enhancing communication.

4) Inclusion of feedback earlier in a "cyber OODA loop" appears to enhance/streamline prediction, which is a potential topic for future research. This could infer that a next step in prediction could involve correction in a manner similar to Kalman filtering. One constraining issue identified is the temporal dimension and the need to process in real-time efficiency.

5) Discerning which capability tier within the DSB framework to characterize an attacker is hard at the beginning of the analysis. For example, methodologies for identifying attackers with capabilities defined in Tiers 1 and 2 (known vulnerabilities–known exploits) are completely deterministic and more precise and defined than attackers in capability Tiers 3, 4, 5, and 6.

6) Related to the previous comment, the IST-129 Task Group notionally agreed that there are unique methodologies for identifying and predicting threats at different levels within the DSB threat capability definition. The implication for practitioners is that for each threat capability family (known vulnerabilities–known exploits; known vulnerabilities–unknown exploits; and unknown vulnerabilities–unknown exploits), this implies each processing stream is different—further supporting the notion that one algorithm does not solve the threat identification or prediction problem.

10.2 Background to the Specialist Meeting

While the growth of available data has grown exponentially, the capabilities of analysis tools, recognition software, and computer capacity has not grown nearly as fast, but these are still much more powerful today than even a decade ago. Several predictive analytic tools that are in the early stages of research show great promise for improving our understanding and ability to support decision-making at reduced levels of risk. At the same time, the challenges of the 21st century have also become more complex and include the impact of a volatile global economy, population migrations, changing weather patterns due to climate change, loss of arable land and fresh water on a global scale, expected population growth, pandemics, and terrorist activities worldwide. Having good indication of likely future actions by nation states, terrorist organizations, refugees, and financial markets has become vital to the planning of collaborative organizations such as NATO to form improved preventative and response strategies to potential large-scale crisis events. The predictive analytic tools available to analysts today are quite powerful when compared to even those of just a decade ago. The problems that can be supported by predictive analysis range from commanding officer decision support in peacekeeping and conflict zones, to strategic decisions based on future global requirements and regional support needs due to predicted pandemic and other health issues, to prediction of natural disasters needing high-availability disaster recovery, to detection of anomalies on critical communication and control data networks in cybersecurity. Some of the required predictions need to be used in decision-making in real time, or even within microseconds of an occurring event, while others can be more strategic and even use massive offline computation. The variables associated with these major challenge areas has led to the development of a collection of predictive analysis tools and research programs with differing properties. There are already a number of tools being developed to provide predictions from the rapidly growing available world databases, but often there is little crosstalk between researchers developing some of the most effective predictive tools.

Approaches exist for the predictive analysis of adversarial COAs in noncyber domains (e.g., Brown et al. 2002; Kott and McEneaney 2006), although the efficacy and robustness of these approaches remains debatable. The shift of military operations to a reliance on cyberspace and the speed of actions in that domain lead to a need to be proactive in understanding how attacks happen and, more importantly, what is likely to occur in the future as a result.

Predictive analysis has been widely relied upon to evaluate options in many domains such as banking, gaming, insurance, and retail. These techniques have not yet been applied to the cyber domain, likely because there are significant challenges in doing so including the following:

- Cyberspace is complex, dynamic, asymmetric, and not well understood, making the adversary’s choice of potential attack steps much larger than in other domains.
- The adversary has the upper hand because their actions in cyberspace are much less observable and take less time than in other domains.
- The rapid evolution of new and unidentified exploits obscures knowledge of the current situation.
- There are diverse cultural, social, and cognitive traits of the adversary that are likely important factors in determining future adversarial COAs.

- Coordination among nations requires close collaboration to enable extremely fast exchange of knowledge about adversaries and their anticipated operations using a common set of concepts, terms, and methodologies.

There are aspects of adversarial action and the cyber domain that can be used to our advantage in predictive analysis. It may be possible to turn the temporal advantage of the adversary's quickness of action to our advantage if we can get inside of their decision-making (OODA) to make timely and accurate predictions of their future actions. We can also use our knowledge of the adversary's capabilities, and the maturity thereof, to reduce the space of possible adversarial actions and increase the accuracy of our predictions.

10.3 Keynote Presentation: Adversary Intent Inferencing for Predictive Analytics

In performing some of the committee analysis in the area of predictive analysis, there was one researcher who had done substantial work in the predictive analytics area. The Specialist Meeting was fortunate to have Dr Santos as the keynote speaker on the topic of "Adversary Intent Inferencing for Predictive Analytics". The focus of the keynote was determining adversary intentions and understanding what drives those actions. The domains of discussion are on military operation, planning, and intelligence analysis.

One reason modeling adversaries is difficult is the level of uncertainty in predictions and the relatively wide open nature of research in this space. Intent inference, or user-intent inference, involves deducing an entity's goals based on observations of that entity's actions (Geddes 1986). In turn, this becomes useful for generation of advice and the definition of future information requirements (Bell et al. 2002; Santos 2003, 2005). There are several approaches to intent inferencing as follows.

Plan-goal-graph (PGG): PGG is a network of plans and goals, where each high level goal is decomposed into a set of plans for achieving it, and the plans are decomposed into subgoals, which in turn are decomposed into lower-level plans (Geddes 1994). Intent is finding the path from observables to a plan or goal.

Operator function model (OFM): OFM is an expert system using a heterarchic–hierarchic network of finite-state automata, in which nodes represent entity's activities and arcs represent conditions that initiate/terminate certain activities and connect observed action to appropriate activity trees (Rubin et al. 1988a; Bushman et al. 1993; Chu et al. 1995).

Generalized plan recognition (GPR): GPR is recognizing the entity's plan for carrying out the task based on observations, an exhaustive set of discrete actions (a plan library), and constraints (Carberry 1988; Lesh et al. 1998; Goodman and Litman 1990).

Intent becomes important because it can help you predict the future, explain the present, and understand the past. Additionally, understanding and identification of intent can help prune the search space, bound optimization, guide scheduling, and better allocate resources.

Traditionally, blue team (i.e., defender) COAs were war-gamed against the "most likely/dangerous" red team (adversary) COAs (circa 2001), but more often prescripted as opposed to being more dynamic. Asymmetry of capabilities and asymmetric threats both mean differences in intent. The question becomes more of an issue of how you do assessments or "what if" analyses.

Essentially the goal is to develop better adversarial modeling. This spawns the question of identifying what you need to know about the adversary. Intent is not just a plan or an enemy COA, but also considers the "why". Some of this can be ascertained by looking at what will happen next. The definition of adversarial intent = Goals + Beliefs + Actions + Commitment. Adversarial modeling becomes useful in financial/business competition (game theory), politics/elections, sports, and so on.

Dr Santos introduced the concept of Dynamic Adversarial Gaming Algorithm (DAGA). DAGA develops algorithmic techniques to accurately predict Community of Interest responses to social, cultural, political, and economic actions. It incorporated various learning aspects; each different play has a different outcome. It gives you a graph of possibilities. Cultural differences were shown to be important with respect to the gaming. What do you need to know about the adversary? What is rational? These questions were based on social, cultural, economic, and political parameters.

It also allows for Bayesian fusion of these factors to model different groups, in different conditions, and make them more asymmetric in simulations. To highlight DAGA's capabilities, it was integrated with the popular Civilization 4 (2005–2008) game engine to demonstrate how the infusion of socio-cultural influences leads to a much-more realistic asymmetric adversary.

Next Dr Santos talked about his most recent work modeling complex adversaries and their intent. This work uses a networked intent model, with evolving behaviors, for multiple adversaries. The goal of this work is to help commanders and decision makers by modeling targets as complex, adaptive systems. The model can produce timely, correct, and actionable intelligence for the Warfighter when the system has only partial observable assets, fluid environments, multi-entity situations with dynamic friends, foes, and neutral parties. He used an example of a Somali pirate group, where the structure of the group was modeled as a network hierarchy with different roles, lines of communication, and social ties.

His future work includes plans for learning adversary intent using dynamic decision models.

10.4 Conclusions and Findings of the IST-145 Specialist Meeting

The IST-145 Specialist Meeting identified several areas where they are researching prediction both within and outside the cyber domain. While some work has been done, not enough in the opinion of the committee; much work still needs to be done in both research and implementation. Our results from this Specialist Meeting are organized into five areas: key results and findings as identified by the committee, some general observations on the practice of prediction, and then some recommendations for the cyber modeling, cyber analytic/algorithm, and cyber prediction communities. Those results are detailed in Section 11.2 where we present all our intermediate- and postSpecialist Meeting findings, observations, and recommendations. The results of the Specialist Meeting are summarized as follows.

Key results and committee findings:

- Several papers introduced multiple algorithmic approaches; for example, one paper described a two-model approach with one checking uniformity of the model with a statistically proven method (Bayesian), whereas the other is checking the autocorrelation (Monte Carlo). Our discussions both during presentations and in the breakout groups concluded that it appears no one algorithm is enough to solve the problem. This appears to support the notion that correlation in cyberspace will not use single-algorithmic approaches.
- The committee felt the specific edge cases of 0% certainty an event will not happen and 100% certainty that an event will happen might be unattainable. This is primarily due to the possibility that events can be independent variables in the computations.
- The committee also concluded that the known vulnerability/known exploit is a solvable problem and has been solved, but not necessarily implemented through automation. It is detection as opposed to prediction, making prediction trivial in this case.

- The committee felt that if you detect trends in the strategic capabilities of an adversary, then this should also be an input to a higher level, overall strategic threat intelligence and prediction system, with respect to this adversary's capability development/improvement and possible new or altered cyber TTPs.

- The committee felt the structure of STIX lends itself to more efficient communications across all entities working the cyber-event prediction problem. STIX already is structured to contain important information and was formed to help security practitioners "to better understand what computer-based attacks they are most likely to see and to anticipate and/or respond to those attacks faster and more effectively".

General observations from the Specialist Meeting:

- Papers mostly addressed analytical approaches with varying degrees of application to the known vulnerability–unknown exploit problem.

- Identifying and understanding a baseline security posture is important to understand the normal state of the network as the initiator to focus on anomalies that deviate from that normal state.

- There is some important research being performed, particularly within the US Economic Development Administration, Defense Advanced Research Projects Agency, and other national research agencies. This work should be monitored and outcomes shared.

- Key cognitive application areas being investigated may include AI for cyber operations, ML for cyber operations, deep learning (neural networks) for cyber operations, human factors for cyber defense, and algorithms' design and engineering.

- Instituting RuNet approaches (closed networks) can adversely affect the ability to do prediction, event correlation, and attribution. The term RuNet refers to a country, such as Russia, isolating its segment of the Internet from the rest of the world (Nikkarila and Ristolainen 2017).

- Most discussions mentioned the lack of valid training data or at least sets of training data where the validity and provenance was certain.

Recommendations to the cyber-modeling community:

- Different state space models (SSMs) require different numbers of samples for operating at the same level of accuracy (even the same SSM at different states). In addition, recent advances in multiple importance sampling and adaptive importance sampling allow using a few samples and still having a great performance.

- Developing a more autonomous intrusion-handling system will require both knowledge, including behavioral, criticality, and impact models, as well as the ability to gain experience (i.e., learning) by leveraging past events.
- Model a closed (national level) network and construct representative cyber-attack scenarios. By doing that, we may be able to extract characteristics of closed network spaces.
- M&S of potential predictions could provide insight into affects and effects of acting on a particular prediction.

Results for the cyber analytics and algorithm community:

- Given the approach from the Bowman paper (Asher et al. 2017), Elizabeth Bowman presented at the Specialist Meeting (that analyzes relationship entities to identify potential members of a threat group), the committee agreed that this approach for this use case is useful in clarifying relationships. Potential application in cyber domain is not so much prediction, but rather given a set of cyber events (the "messages" from this paper) what could be hierarchy or the relationship across those events.
- Developing attack graphs around known vulnerabilities could generate all, or most all, of the possible attack paths. This approach may be able to reduce the prediction problem (for the known, unknown case only) to a more deterministic approach that concentrates on likelihood of a graph event occurring.
- Anomaly- and signature-based detection inputs can be combined based on an analysis of past results of event logs.
- Some discussion pointed out that if an adversary compromised the predictive analytics, that adversary could manipulate inputs exploiting the algorithm and corrupting results.

Results addressing cyber prediction:

- Although some of the research is novel and interesting, planning is not prediction. Planning is analytical and partial mathematical approach whereas prediction results are better served via a mathematical approach.

- Using the known vulnerabilities as a mechanism to produce attack graphs identifying potential exploits can reduce the space of uncertainty in predictions.
- When talking about predictions, anything is possible (within certain universal limitations). Predictions in cyberspace are not limited to certain physical or temporal constraints. Cyberspace does not have traditional physical constraints. Because (almost) anything is possible in the future in cyberspace, the space of possible (if unlikely) outcomes is extremely large. Therefore, this space is difficult to model and simulate.
- Discussions indicated we may not be able to predict with certainty, but we may be able to predict likelihood.
- A predictive system could be applied to other areas of cyber defense to potentially help prioritize future patching and allocation of defensive resources including identification of adversarial deception and use the predictive analysis to select potential COAs.

11. Technical Activities Recommendations

During the course of the research, the team identified some obstacles that form the basis of our recommendations. In short, we pinpoint three areas that warrant attention from the research community. First, there is a lack of completed or envisioned empirical research in true prediction. And by that, we mean "what will the next attack look like", not what is the *likelihood* of an attack or the *likelihood* of the next event. Within cyberspace, the next event may or may not be dependent on the previous attack(s) or the dependency may not be understood at this time by the defenders. Second, we saw a constant and consistent observation in the cyber field and that is the lack of completed or envisioned empirical research, experimentation, or development of realistic data that could be used in predictive research. The lack of realistic data sets has forced many analytics or algorithms to be mathematically proven, possibly through formal methods. While formal methods provide substantive verification and proofing, the concept does not ensure an algorithm or analytic is correct. Rather, the methodology highlights errors related to "inconsistencies, ambiguities, and incompleteness that might otherwise go undetected" (Clarke and Wing 1996).

Typically, NATO RTGs recommend continuations in their specific research areas. In the case of IST-129, we have several reasons for not doing this *exactly*. First of all, we are recommending both an experiment evaluating prediction using combinatorial analysis and an evaluation of expanding attack surface analysis for the known vulnerability–known exploit adversary, reducing that part of the problem to a deterministic one that is aligned with our findings. Second, as of the writing of this report in mid-2019, there are several other TAPs being proposed by the NATO technical community where we believe cooperation with those groups will yield results. The two TAPs being proposed are an ET for unsupervised ML in military domain and leveraging cyber-range capabilities for security M&S in support of secure system development and security certification testing where we believe cooperation with that RTG will yield results. Working with the unsupervised learning ET, we would expect to further information and knowledge on potential algorithms and approaches that might impact the prediction problem. The cyber range directly addresses one of our recommendations on the development of more realistic data sets that would be useful to the research community working the prediction problem. The lack of a data set that represents real traffic in either commercial or military environments not only hampers solution testing, but prevents benchmarking proposed solution effectiveness.

Additionally, when creating comparative methodologies for predicting the potential for cyber-attacks, we believe the tools financial institutions and credit scoring companies have created and used to predict the repayment of loans could provide some additional insight, even though in their case this is a bounded problem. One direction here seems to be the combination of similarity coefficients, logistic regression, ML tools, and Bayesian networks for forecasting situations (loan repayment/nonreturn). The appropriate ways and means of attending seem to be applicable to the prediction of things in the frame of cyber-attacks.

There appears to be promise using predictive statistics with randomized and anatomized experiments. This would include experiments that used general independent variables like human error (H_E), human assumptions (H_A), organizational error (O_E), technical assumption (T_A), and technical error (T_E). While we recognize this could be tangentially related to the predictability problem, we did not make a recommendation in this area.

11.1 Specific RTG Recommendations

RTGs based their efforts on predicting adversarial behavior that is centered on attack-surface analysis expansion. Another observation from IST-129 was that predicting behavior for threats using known vulnerabilities with unknown exploits could be accomplished by expanding attack trees, which reduces the problem set from one of prediction to one of pattern matching. The intent of this RTG would be to develop those approaches to expand attack trees around known vulnerabilities to encompass unknown exploits.

One recurring recommendation is in an area that is beyond predictability, but is an important task needed to support effective research on predictability—and that is the data set deficiency for evaluation of any potential solution. This data set has to be realistic and yet contain some undiscovered patterns that represent unknown exploits and then also a representative set of noise. There is, as of 2019, a pending RTG for Leveraging Cyber-Range Capabilities for Security Modeling and Simulations in Support of Secure System Development and Security Certification Testing. IST-129 is in full support of this recommendation and will support the formal establishment of that RTG. As such, we will not submit a competing RTG, but rather work within this recommendation to begin development of the data sets. The following represents a list of issues IST-129 would like to see addressed within the cyber-range community or the M&S community. The specific issues for study are the following:

1) Using current approaches and prediction algorithms, which approach is best suited for developing a predictive capability in cyberspace, and which level of maturity can currently be expected? Some domains have limitations on the predictive model. For example, a conventional weapon model for a missile would not have to consider a case where a launched missile could "teleport" from a northern trajectory to a western one instantaneously. Cyberspace can have attacks "appear" from many locations simultaneously. We believe that not all methods of prediction can be ported to the cyberspace domain, so testing is required to support or refute this assertion. More specifically, which theoretical model (existing, combination of existing, or to be developed) is best suited to evaluate and compare the effectiveness of different predictive techniques in a military decision-making context?

2) Cyberspace as a domain has several considerations that make it different from other domains. The nature of cyberspace, the speed at which attacks unfold, the lack of traditional physical constraints, and so on, imposes specific constraints on an adversary behavior-prediction approach, such as high-speed prediction for immediate tactical/technical reaction and lower speed for long-term strategic prediction. To what extent do current techniques match these constraints? Again, this is an algorithmic implementation issue and testing is needed to provide some empirical evidence on which to construct or base solutions.

3) What is the role of the human-in-the-loop in the process of predicting adversary behavior in cyberspace; which visualizations, symbology, and so on, still need to be developed to optimize their effectiveness? While the original remit was to uncover automated approaches to prediction, current cyber-defense approaches have human-in-the-loop. There is potentially a great deal to be gained from fusing predictability analytics with visualization techniques to force-multiply the human-in-the-loop.

11.2 Immediate ET Recommendation: ETs on Closed Networks

IST-129 strongly recommends an ET focusing on closed networks. Closed national segments of the internet may be formed in the near future. For example, Russia has declared its aim to become "digitally sovereign". One part of "digital sovereignty" is the capability to close off its national segment from the global internet whenever required and maintaining operational capabilities of the national segment while doing so. This directly addresses the challenge to NATO in recognizing threats and precursor behavior from disruptive technologies that exist behind closed networks. So we view this as an opportunity to predict how situational awareness is altered via the establishment of closed networks.

A Cyber Defense Situational Awareness (CDSA) capability is an emerging, urgent need across nations. CDSA plays a vital part in this requirement. In addition, several nations are developing CDSA tools, techniques, and technologies, and are at the point where they could leverage each other's efforts through international collaboration. These factors indicate a timely opportunity for international collaboration, thus it is recommended to create a new ET that addresses related research and technological issues in the area of CDSA.

Effective CDSA requires the integration of multiple components' situation awareness, potential implications, and COAs. The exploratory group should review the state of the art and address these focus areas.

- Situation awareness: What is the current state of the art related to relevant closed national segments of the internet? In addition, which countries are planning or already implementing closed national segments?

- Potential implications: What is the potential implication to cyber defense for the NATO alliance? Also, risk assessment is required for decision making and to achieve mission assurance. This risk assessment must be dynamic, that is, produced from the continuous monitoring of the cyber environment and not from the traditional static threat and risk assessment approach.

- COA: To determine what could be a proper COA for NATO, at the technical level, or to respond to the challenge introduced by the formation of closed national segments of the internet. To determine COAs, it is necessary to resolve, analyse, and understand different alternatives' implications.

11.3 Other ET Recommendations

One observation of the IST-129 team was the solution that tactical use of ETs may in the long run be more beneficial in sparking development for certain problem areas that are still not clearly defined. The solution development or even approaches for both the threats using unknown exploits against known vulnerabilities and the threats using unknown exploits against unknown vulnerabilities is challenging due to the rate of change in new and unknown attacks appearing. In addition, cyber defense as a science cannot wait three or more years for a typical RTG to complete research results. It may in the end be better to have partial results that can be used to develop more effective cyber defenses. Because of the increasing velocity and variety of cyber-attacks that cyber-attacks change, IST-129 is suggesting the formation of two tactical ETs:

1) The first recommended ET would evaluate the viability of using known TTPs as a means of predicting adversarial behavior. One of the observations made by IST-129 was that adversarial behavior could be evaluated by examination of known TTPs and then extrapolating after so many events to predict the next event, making the assumption that events are linearly dependent. This thought is not without caution, since another observation made is that next events could also be thought of as linearly independent. The concept has some level of usefulness and IST-129 believes an ET could provide some scientific support for/against such an approach.

2) The second recommended ET would investigate identifying adversarial behavior based on attack tree and attack surface analysis. This approach is based on the notion that for known vulnerabilities it could be possible to develop all possible attack exploits thereby eliminating unknown exploits and reducing the problem of predictability into one of pattern matching. This approach works only for the adversaries leveraging known vulnerabilities with either known or unknown exploits. The salient feature of this approach is that it reduces the overall problem space of predictability of adversarial behavior to finding methods to address the adversary who uses/develops unknown vulnerabilities with unknown exploits. This recommendation is consistent with the committee's findings that addressing

adversarial behavior and predicting “next moves” will be a multimethod solution.

11.4 Specific Research Projects and Experiments

The committee saw the need for specific testing that could be conducted by one or more NATO member countries or as an exercise of cyber capability. Specifically was the need to use real and known adversary TTPs against a NATO or member environment. We recognize this could be a classified exercise, so we would anticipate this would be a limited experiment and one that could almost be a tabletop red team approach, but that the results of such an experiment could provide deeper insight into *how* to use predictability approaches. This could also be approached through an ET that would design and execute an experiment using combinatorial analytics for predictability. One of the observations made by IST-129 was that any analytic or process that could be used to predict adversarial behavior will be a multi-algorithmic method. It will more likely be a compound analytic that is more OODA-loop organized. This activity would evaluate candidate compound architectures and provide some tangible results. Analytics for prediction need to become standard components of the cyber-gaming process and for the present, must be simulated in exercises to illustrate how to effectively use prediction technology as it matures to eventually uncover, detect, and prove adversary attacks.

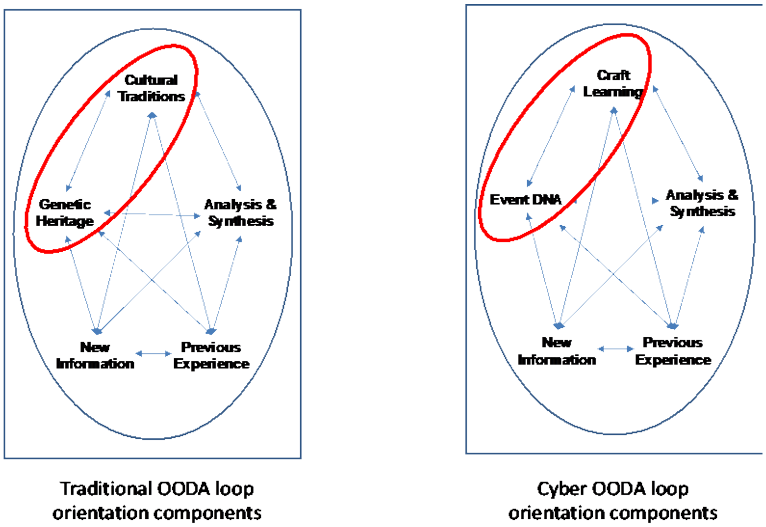
11.5 Recommended Areas for Research and/or Position Papers

We noted during the literatures searches there were some areas that appeared to be under-researched where we hoped to find some useful information. We proffer several questions and topics for academic research and consideration.

Are there specific laws of physics in cyberspace? This arose from noting that prediction of movement (e.g., in a fighter aircraft) can be tracked via radar with accurate movement predictions using Kalman filtering. Part of the reason for the ability to accurately track stems from the laws of physics that prevent the aircraft from instantaneously changing position or making a 120° reverse turn. In cyberspace, these laws do not apply. An attack that follows a pattern can instantaneously change that pattern or show up in a completely different location. Research into the equivalent constraints in the cyber world to describe freedom of movement in cyberspace, what controls/restricts it, how/what actions are available, if events are both independent and dependent at the same time, how actions are restricted by environment, and so on, could be of extreme benefit to cyber-defense practitioners.

Develop a primer for predictive analytic implementation methods. The IST-129 group expected with the amount of research and practice into analytic engines that information existed on which type of algorithmic implementation worked (or did not work) for a specific problem set or environment. For example, what domains would be suited or underserved for implementing unsupervised learning and what domains are better suited. The output of this research would be a matrix of analytic approach (implementation) cross-referenced against environment of use and specific implementations. This specific research omission was noted by the committee as a necessary part of a comprehensive approach to predictability. Furthermore, this would provide a much-needed reference guide that could be applied to problems of predictability and the best predictive models in different domains.

The Cyber OODA Loop. Part of our study used the OODA loop as it could be applied to the problem of adversarial cyber predictability. The OODA loop is well grounded in the physical world, but we saw little in the way of consistency or usability in using a cyber version of OODA in the cyber world. We spent the better part of one of our meetings discussing this issue and reviewing individual cyber OODA papers. Specifically, the area for study would be the orientation portion of the OODA loop, as illustrated in Fig. 9. Event DNA and how/where an adversary learned the craft could have a large impact on the analytic bias—in a good way. We believe this could be used as an ordering scheme for identifying the most-likely to least-likely prediction.



Traditional OODA loop orientation components

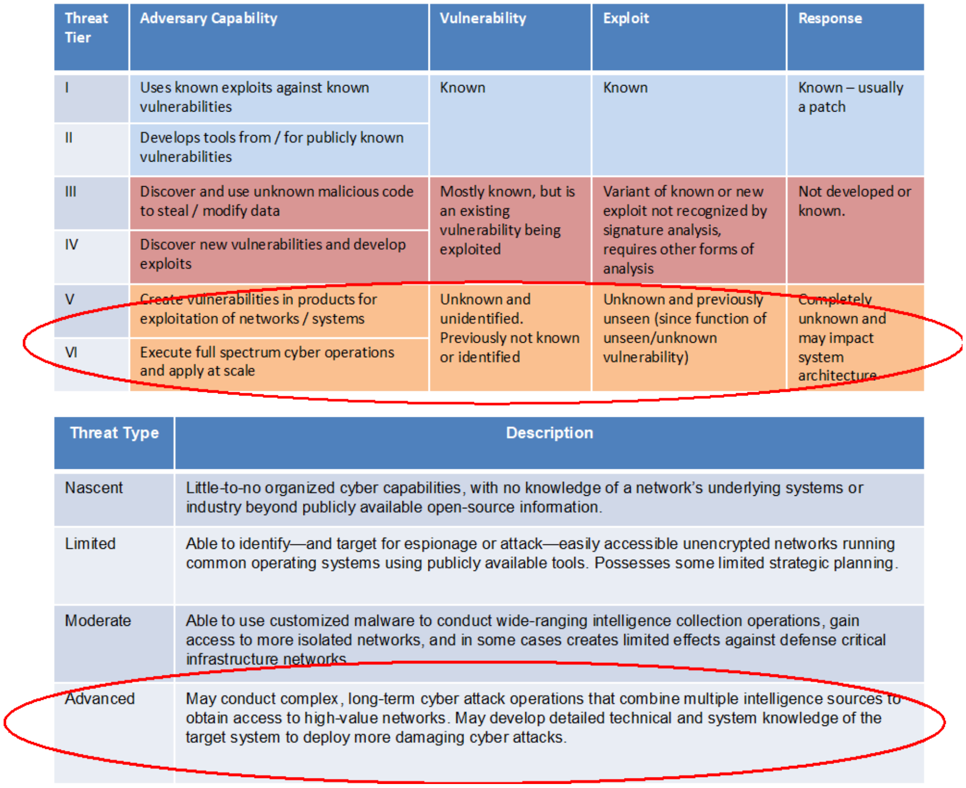
Cyber OODA loop orientation components

Fig. 9 Orientation phase of OODA loop with cyber unique area highlighted

Taxonomy for predictive cyber analytics. A small but relevant research activity could be to recommend a taxonomy for predictive cyber analysis. For describing the threat, the RTG preferred to use a capability-based approach, as shown in Tables 1 and 2. While these descriptors do not include TTPs, they do reference the inherent capability and by abstraction, the knowledge base of the threat. The RTG used STIX and TAXII as vehicles for structuring our thinking. While the remit of the IST-129 did not include the development of a predictive cyber taxonomy, that would be an activity that could help provide structure for consistent and defined communication on and categorizing of predictive analytics.

During discussions across the RTG, we identified some issues that we put in a "parking lot" because while they were important and potentially related to the predictability problem, they were outside the remit of the group. We believe that these should be discussed and addressed further:

a) There is a concern with threats who manage to evade primary ingress detection. Postdetection researching capabilities are necessary when a threat bypasses defenses or uses new facts to enter a network to minimize damage. Some focus should be given to those threats since they might exhibit discernible behavior that could be identified in postdetection system processing. This represents an entirely new approach to threat detection as this will be inside the system rather than at the perimeter. This threat would be expected to be the highest-capability threat, as illustrated in Fig. 10. This is a far-more risky approach, since the threat would have already established residence in the system but could represent an additional means of layered cyber defense.



Threat Tier	Adversary Capability	Vulnerability	Exploit	Response
I	Uses known exploits against known vulnerabilities	Known	Known	Known – usually a patch
II	Develops tools from / for publicly known vulnerabilities			
III	Discover and use unknown malicious code to steal / modify data	Mostly known, but is an existing vulnerability being exploited	Variant of known or new exploit not recognized by signature analysis, requires other forms of analysis	Not developed or known.
IV	Discover new vulnerabilities and develop exploits			
V	Create vulnerabilities in products for exploitation of networks / systems	Unknown and unidentified. Previously not known or identified	Unknown and previously unseen (since function of unseen/unknown vulnerability)	Completely unknown and may impact system architecture
VI	Execute full spectrum cyber operations and apply at scale			

Threat Type	Description
Nascent	Little-to-no organized cyber capabilities, with no knowledge of a network's underlying systems or industry beyond publicly available open-source information.
Limited	Able to identify—and target for espionage or attack—easily accessible unencrypted networks running common operating systems using publicly available tools. Possesses some limited strategic planning.
Moderate	Able to use customized malware to conduct wide-ranging intelligence collection operations, gain access to more isolated networks, and in some cases creates limited effects against defense critical infrastructure networks.
Advanced	May conduct complex, long-term cyber attack operations that combine multiple intelligence sources to obtain access to high-value networks. May develop detailed technical and system knowledge of the target system to deploy more damaging cyber attacks.

Fig. 10 Specific areas of threat capability that may be suited for a post-ingress detection capability

b) There continues to be calls for increasing the efficacy of cyber-threat modeling. An additional requirement for threat models would be to ensure the evolution of cybersecurity threats models, techniques, and tools to account for changing adversary behavior. Threats continue to evolve new approaches and toolsets, all of which are necessary to understand how networks are going to be compromised by an advanced capable threat.

c) Developing and sharing real-threat source data continues to be problematic with privacy and proprietary restrictions. However, evaluation of new analytics and algorithms to identify and detect evolving threats would benefit from standardized benchmarking data. Of extreme interest would be approaches that could develop and share benchmarking data and maintain currency of that data.

d) Cyber adversarial-prediction M&S. The nature of cyberspace, the speed at which attacks unfold, the lack of traditional physical constraints, and so on, imposes specific constraints on adversary behavior-prediction approaches, such as high-speed prediction for immediate tactical/ technical reaction and lower speed for long-term strategic prediction. To what extent do current techniques match these constraints? Secondly, what is the role of the human-in-the-loop in the process of predicting adversary behavior in cyberspace? Which visualizations, symbology, and so on, still need to be developed to optimize this effectiveness? And finally, which theoretical model (existing, combination of existing, or to be developed) is best suited to evaluate and compare the effectiveness of different predictive techniques in a military decision-making context?

11.6 Parameters on Experiments Recommendation

The RTG did not develop any specific experiments as part of this final report, but notes some considerations for groups formulating experiments as these may provide additional information toward a cyber predictability solution.

Situation awareness and understanding is crucial in the cyber domain. As noted earlier, it is essential to intensify and deepen the modeling of national segments of the internet. It is as, or even more, important to develop experiments and exercises related to the closed national networks to extract their factual effects in cyberspace. One should design and execute experiments from technical up to strategic levels. The technical-level experiments' results may be used as an input for strategic-level experiments. Strategic-level experiments can be also conducted separately, for example, by applying TTXs (Lantto et al. 2019). A TTX can be organized more quickly and it could assist in creating the situation awareness of the implications of the closed national segments of the internet. Knowing unfavorable COAs in advance is highly advantageous, and furthermore, to know a closure's potential effects helps avoid adversary deception. Consequences of closing national networks can be studied by using matrix wargaming methods to convey the complex interdependencies and interactions (Lantto et al. 2018). Even though one would not consider closing alliances' national segments of the internet as a response to the potential adversaries' closed national segments, it would be beneficial to analyze the effect of such closure. A closure can be unintentional as well.

It is acknowledged that the amount parameters are statistically significant. Consequently, prior organizing an exploratory experiment of any kind of the variables to be tested in the experiment have to be chosen from an actual environment. So we need to set hypothesis on which to base the experiment on group A and group B (blue teams and the red team) on a real NATO exercise. These variables we have to test to explore the relationships, which includes correlation and regression.

Experiments alluded to in the literature on a cyber-range were run to measure temporal and accuracy constraints, evaluate scalability, COAs, usability of visualization for decision support, and protection mechanisms for cyber-defensive suites.

12. Conclusions and Ultimate Findings

While there are interim findings and results of the committee described throughout this report, here we gather together an overall summary.

1) **Prediction of adversarial operations in cyberspace is complex, but can be decomposed.** The prediction problem of adversarial operations within cyberspace is not 100% solved, and although events may be connected, they may not be treated as dependent variables. The committee concluded that the problem can be reduced in several ways:

    a) Although now we may not be able to predict adversarial operations with certainty, we may be able to predict likelihood (plausibility) of the next adversarial operation.

    b) A predictive system could be applied to other areas of cyber defense to potentially help prioritize future patching and allocation of defensive resources including identification of adversarial deception and use the predictive analysis to select potential COAs.

    c) Attack graphs (developing likelihood similarities to identified and known exploits) are one known methodology identifying exploits leveraging known vulnerabilities and can reduce the space of uncertainty in predictions of potential exploits.

2) **Predicting adversarial operations will be a multi-algorithmic approach and not a singular methodology.** Contrary to the conventional thinking of singular method algorithms to identify all threats in cyberspace, correlation in cyberspace will not use single algorithmic approaches. The DSB report (Gosler and Von Thaer 2013) indicates three capability levels of threat and the committee found evidence that identifying and predicting adversarial operations will be a multi-algorithmic approach. This implies that single algorithmic approaches do not work for prediction against different capability threats.

    a) The committee concluded that the detection problem of a known vulnerability/known exploit (classic statistical approaches) is a solvable problem and has been solved, but not necessarily implemented through automation. It is detection as opposed to prediction, leaving prediction in this case still a challenge.

b) The IST-145 Specialist Meeting presented information that algorithms and processes to identify and possibly predict threats in the (known, unknown) region of the DSB definition will be multimethod ensembles as opposed to single-method algorithms. This goes against conventional thinking of singular-method algorithms to identify all threats in cyberspace.

c) The threat with (unknown, unknown) capabilities may be an area better suited for AI algorithmic approaches. Developing data sets for the (unknown, unknown) threat will be complex and challenging, impacting the accuracy of learning algorithms due to smaller and incomplete data sets and small descriptions of situations and processes.

3) **STIX is a facilitator for cyber prediction.** The committee felt that the structure of STIX lends itself to more efficient communications across all entities working the cyber-event prediction problem. STIX is already is structured to contain important information and was formed to help security practitioners better understand what computer-based attacks they are most likely to see and to anticipate and/or respond to those attacks faster and more effectively.

4) **Cyber defense itself needs to be protected**. The committee noted that prediction approaches need to be made resistant to cyber deception or manipulation, which is critical to maintain trusted operations. Cyber-defense decisions and COAs based on untrusted predictions will reduce the ability to maintain resilient operations. This area certainly needs further research.

5) **Modeling of closed network systems is needed.** The goal of prediction of adversarial operations within closed networks is more complex because of a lack of data. This will require models of closed networks, realistic closed network testing data, attack scenarios, and development of new cyber-defense paradigms. New attack vectors for a closed and controlled internet border need to be found and defined. Command and control structures used within and against a closed and controlled internet border are also not well defined and could be difficult to discover.

6) **We need data sets that are representative of reality.** There is a lack of clear benchmarks with respect to data sets, both in terms of existence or being realistic and representative of actual network traffic. This includes both actual attack data, hidden data, obscured data, and noise that would provide agreed upon data sets for benchmarking. This implies an inability to compare solutions for how things should perform due to a lack of agreed measurements and assessments that represent ground truth.

7) **Threat level and attacker capability must be understood.** For actionable predictions, we will need to identify the level of threat posed and attacker capabilities; this is much easier to manage for the lower-capability tiers and not yet understood for higher levels.

8) **Observe changes in strategic capability.** Trends in the strategic capabilities of an adversary should also be an input to a higher-level threat intelligence and prediction system for possible new or altered cyber TTPs.

9) **Notice similarity and trajectory.** One way to predict what happens next in cyberspace and what is its credibility level is to rely on numerical evaluation of similarity for situations and developments. Knowledge of where an adversary exists in physical reality and in cyberspace, and its trajectory, is important to take into account.

10) **Implement speedy predictions.** To be effective, prediction must be swift enough to enable timely response, which is contrary to current state of the practice. An autonomous and predictive system must consider both short-term predictions and longer-term trends.

11) **Prediction should fit into a feedback loop.** For example, feedback earlier in a "cyber OODA loop" might help to iteratively refine results (subject to speed requirements).

12) **More data are needed.** Data used for strategic intelligence that consists of values put on core assets are particularly needed, as well as catalogues of known or anticipated threat actors and their TTPs. Historical attack data is essential to estimate risk and make predictions.

13) **Results cannot be certain.** No software or hardware tool for prediction analysis is exact and so errors must be managed. Research is needed to quantify errors and understand uncertainty ranges possible in predictive outputs and it is important to distinguish between what is known absolutely from what is known only with some degree of uncertainty.

14) **Follow data protection and compliance.** The legal and regulatory landscape is rapidly changing in response to widespread use of analytics and targeted social media. Predictive analytics systems that use inputs from personal or personnel resources must be aware of their use of sensitive data or potential for biased results.

15) **Existing commercial products.** While not a recommended or complete product list, the committee at various times referred to some of the commercially available products to help identify what data could be available,

the format of the data, and any processing that currently exists: Norse, Checkpoint SW, FireEye, Arbor Networks, Trend Micro, Akamai, Fortinet, Splunk, and ArcSight.

13. References

Amazon Web Services. NIST cybersecurity framework (CSF). Aligning to the NIST CSF in the AWS cloud; 2019 Jan. https://d0.awsstatic.com/whitepapers/compliance/NIST_Cybersecurity_Framework_CSF.pdf.

Angerman WS. Coming full circle with Boyd's OODA loop ideas: an analysis of innovation diffusion and evolution [theses and dissertation]. Air Force Institute of Technology; 2004 Mar 12. https://scholar.afit.edu/cgi/viewcontent.cgi?article=5087&context=etd.

Asher D, Caylor J, Mittrick M, Richardson J, Heilman E, Bowman E, Korniss G, Szymanski B. The investigation of social media data thresholds for opinion formation. In: Proceedings of the 22nd international Command and Control Research & Technology Symposium; 2017 Nov; Los Angeles, CA. https://arxiv.org/abs/1712.04100.

Ashfaq RA, Wang XZ, Huang JZ, Abbas H, He YL. Fuzziness based semi-supervised learning approach for intrusion detection system. Inf Sci. 2017;378(C):484–497.

Barnum S. Standardizing cyber threat intelligence information with the Structured Threat Information eXpression (STIX). Mitre Corporation. 2012;11:1–22.

Bell B, Santos Jr E, Brown SM. Making adversary decision modeling tractable with intent inference and information fusion. In: Proceedings of the 11th Conference on Computer Generated Forces and Behavioral Representation; 2002 May; Orlando, FL.

Boyd JR. Patterns of conflict. Unpublished paper; 1986 Dec.

Brown SM, Santos E Jr, Bell B. Knowledge acquisition for adversary course of action prediction models. In: Proceedings of the AAAI 2002 Fall Symposium on Intent Inference for Users, Teams, and Adversaries; 2002.

Brumley L, Kopp C, Korb K. The orientation step of the OODA loop and information warfare. In: Proceedings of the 7th Australian Information Warfare and Security Conference; 2006 Dec; Perth, Australia. p. 20.

Bushman JB, Mitchell CM, Jones PM, Rubin KS. ALLY: an operator’s associate for cooperative supervisory control systems. IEEE Trans Sys Man Cybernetics. 1993;23(1):111–128.

Buczak A, Guven E. A survey of data mining and machine learning methods for cyber security instruction detection. IEEE Commun Surv Tutor. 2016;18(2)1153–1176. doi:10.1109/COMST.2015.2494502.

Carberry S. Modeling the user's plans and goals. Comput Linguist. 1988;14(3):23–37.

Chu RW, Mitchell CM, Jones PM. Using the operator function model and OFMspert as the basis for an intelligent tutoring system: towards a tutor/aid paradigm for operators of supervisory control systems. IEEE Trans Sys Man Cybernetics. 1995;25(7):1054–1075.

Clarke EM, Wing JM. Formal methods: state of the art and future directions. ACM Computing Surveys (CSUR). 1996;28(4):626–643.

Connolly J, Davidson M, Schmidt C. The trusted automated exchange of indicator information (TAXII). The MITRE Corporation. 2014;1–20.

[DOD] Joint publication 2-0, joint intelligence. Department of Defense (US); 2013 Oct 22.

[Galois]. Framework for information disclosure with ethical security (FIDES). Galois, Inc. [accessed 2019 Apr]. https://galois.com/project/fides/.

[GAO]. Weapon systems cybersecurity: DOD just beginning to grapple with scale of vulnerabilities. Government Accountability Office (US); 2018 Oct. Report No.: GAO-19-128. https://www.gao.gov/assets/700/694913.pdf.

Geddes N. The use of individual differences in inferring human operator intentions. AAAIC ’86. Aerospace Applications of Artificial Intelligence. IEEE; 1986:31–41.

Geddes ND. A model for intent interpretation for multiple agents with conflicts. In: Proceedings of IEEE International Conference on Systems, Man and Cybernetics; 1994 Oct 2–5; San Antonio, TX. IEEE;1994;3:2080–2085. https://ieeexplore.ieee.org/abstract/document/400170/.

Goodman BA, Litman DJ. Plan recognition for intelligent interfaces. In: Proceedings of the Sixth Conference on Artificial Intelligence Applications. IEEE; 1990:297–303.

Gosler JR, Von Thaer L. Task force report: resilient military systems and the advanced cyber threat. Defense Science Board, Department of Defense (US). 2013;41.

Kolter JZ, Maloof MA. Learning to detect and classify malicious executables in the wild. J Mach Learn Res. 2006;7(12):2721–2744. https://www.jmlr.org/papers/volume7/kolter06a/kolter06a.pdf.

Kott AW, McEneaney WM, editors. Adversarial reasoning: computational approaches to reading the opponent's mind. CRC Press; 2006.

Krekel B. Capability of the People's Republic of China (PRC) to conduct cyber warfare and computer network exploitation. Northrop Grumman Corp; 2009 Oct.

Krekel B, Adams P, Bakos G. Occupying the information high ground: Chinese capabilities for computer network operations and cyber espionage. Int J Comput Res. 2014;21(4):333.

Kukkola J, Ristolainen M, Nikkarila J-P, editors. Game changer: structural transformation of cyberspace. Publication 10. Finnish Defence Research Agency; 2017.

Kukkola J, Ristolainen M, Nikkarila J-P, editors. Game player: facing the structural transformation of cyberspace. Publications 11. Finnish Defence Research Agency; 2019.

Lamolinara V. Cybersecurity as it applies to the survivability key performance parameter [presentation]. Defense Acquisition University, Mid-Atlantic Region; 2018 June 6. https://www.dau.mil/Lists/Events/Attachments/104/06-06-2018_Cyber%20Survivability% 20Webinar %20 Final%20with%20resources%20UTM.pdf.

Lantto H, Åkesson B, Kukkola J, Nikkarila J-P, Ristolainen M, Tuukkanen T. Wargaming a closed national network: what are you willing to sacrifice? In: MILCOM 2018; 2018 Oct 29–31; Los Angeles, CA.

Lantto H, Åkesson B, Suojanen M, Tuukkanen T, Huopio S, Nikkarila J-P, Ristolainen M. Wargaming the cyber resilience of structurally and technologically different networks. Sec Def Q. 2019;24(2):51–64.

Lesh N, Martin N, Allen J. Improving big plans. In: AAAI/IAAI '98. Proceedings of the Fifteenth National/Tenth Conference on Artificial Intelligence/Innovative Applications of Artificial Intelligence; 1998 July. p. 860–867.

McCallam D, Braun T, Faganel R, Guenther H, LoPiccolo J, Lorents P, Mees W, Nikkarila J-P, Sommestad T, Varga M. Findings and excerpts from the North

Atlantic Treaty Organization (NATO) specialist meeting IST-145, predictive analytics and analysis in the cyber domain. 2018 Dec. Unpublished.

McCallam D, Braun T, Wunder M, Santos E, Sommestad T, Arregi VE, Bugallo M, Bonneau R, Bowman E, Mittrick M, et al. Approaches to prediction of cyber events: report of the 2017 specialist meeting by the North Atlantic Treaty Organization (NATO) research group IST-145-RTG. DEVCOM Army Research Laboratory (US); 2019 June. Report No.: ARL-SR-0418.

McCallam D, Akesson B, Aspinall D, Braun T, Faganel R, Guenther H, Kellet M, LoPiccolo J, Lorents P, Mees W, Nikkarila J-P, Sommestad T, Varga M. Final report and recommendations of the North Atlantic Treaty Organization (NATO) Research Task Group IST-129, predictive analysis of adversarial cyber behavior. NATO Collaboration Support Office; 2021.

[NATO] North Atlantic Treaty Organization Science and Technology Organization. Information Systems Technology activity proposal for IST-129 Research Task Group—predictive analysis of adversarial cyber behavior; 2015. https://www.sto.nato.int/_layouts/listform.aspx?PageType=4&ListId={B2AAC100-BE82-43CE-886F-0A467BD6BAA9}&ID=16546.

Nikkarila JP, Ristolainen M. 'RuNet 2020'– deploying traditional elements of combat power in cyberspace? In: Proceedings of 2017 International Conference on Military Communications and Information Systems (ICMCIS); 2017 May 15–16; Oulu, Finland. https://ieeexplore.ieee.org/iel7/7950754/7956469/07956478.pdf. IEEE; c2017. p. 1–8.

[NIST] SAMATE – software assurance metrics and tool evaluation. c2018 [accessed 2021 Feb 1]. https://samate.nist.gov.

Pendlebury F, Pierazzi F, Jordaney R, Kinder J, Cavallaro L. TESSERACT: eliminating experimental bias in malware classification across space and time. USENIX Security '19. 28th USENIX Security Symposium; 2019 Aug 14–16; Santa Clara, CA. CoRR abs/1807.07838. http://arxiv.org/abs/1807.07838. p. 729–746.

Perdisci R, Lee W, Feamster N. Behavioral clustering of HTTP-based malware and signature generation using malicious network traces. In: NSDI'10: Proceedings of the 7th USENIX Conference on Networked Systems Design and Implementation; 2010 Apr 28. USENIX Association; Vol 10, p. 14. https://www.usenix.org/legacy/event/nsdi10/tech/full_papers/perdisci.pdf.

Rubin KS, Jones PM, Mitchell CM. OFMspert: inference of operator intentions in supervisory control using a blackboard architecture. IEEE Trans Sys Man Cybernetics. 1988a;18(4):618–637.

Rubin RB, Perse EM, Barbato CA. Conceptualization and measurement of interpersonal communication motives. Hum Commun Res. 1988b;14(4):602–628.

Santos E. A cognitive architecture for adversary intent inferencing: structure of knowledge and computation. In: Enabling Technologies for Simulation Science VII International Society for Optics and Photonics. 2003;5091:182–194.

Santos E Jr. Adversarial intent inference for predictive battlespace awareness. Air Force Research Laboratory (US); 2005 Nov. Report No.: AFRL-IF-RS-TR-2005-378. https://pdfs.semanticscholar.org/56c8/e179e361c907231027a26a989fe4d585b565.pdf#page=8.

Shadowserver Foundation [accessed 2019 Apr]. https://shadowserver.org/.

Skierka I, Morgus R, Hohmann M, Maure T. CSIRT basics for policy-makers. Global Public Policy Institute; 2015 May.

Sorensen CL. Cyber OODA: a candidate model for cyberspace engagement. Air Univ Maxwell AFB AL School of Advanced Airpower Studies; 2010.

Yuen J, Turnbull B, Hernandez J. Visual analytics for cyber red teaming. VizSec 2015. 2015 IEEE Symposium on Visualization for Cyber Security; 2015 Nov. https://doi.org/10.1109/VIZSEC.2015.7312765.

Zhu M, Hu Z, Liu P. Reinforcement learning algorithms for adaptive cyber defense against heartbleed. In: Proceedings of the First ACM Workshop on Moving Target Defense; 2014 Nov 3. https://doi.org/10.1145/2663474.2663481. p. 51–58.

List of Symbols, Abbreviations, and Acronyms

AD	Active Directory
AI	artificial intelligence
APT	advanced persistent threat
AWS	Amazon Web Services
BGP	Border Gateway Protocol
Cap	capabilities
CD	compact disc
CDSA	Cyber Defense Situational Awareness
CIO	chief information officer
CISO	chief information security officer
COAs	courses of action
CSF	Cybersecurity Framework
CSIRTs	Computer Security Incident Response Teams
CTI	cyber threat intelligence
DAGA	Dynamic Adversarial Gaming Algorithm
DEVCOM	US Army Combat Capabilities Development Command
DNA	deoxyribonucleic acid
DNS	Domain Name System
DOD	Department of Defense
DSB	Defense Science Board
ET	exploratory team
FIDES	Framework for Information Disclosure with Ethical Security
FKIE	Fraunhofer Institute for Communication, Information Processing and Ergonomics
GAO	United States Government Accountability Office
GPR	generalized plan recognition

HoI	holder
I&W	indications and warnings
IDS	Intelligent Data and Security
IMPACT	Information Marketplace for Policy and Analysis of Cyber-risk & Trust
Ind	inducer
IP	Internet Protocol
IST	Information Systems Technology
IT	information technology
M&S	modeling and simulation
ML	machine learning
NATO	North Atlantic Treaty Organization
NIST	National Institute of Standards and Technology
NU	NATO Unclassified
OFM	operator function model
OODA	Observe, Orient, Decide, Act
Opt	option
PGG	plan-goal-graph
Res	resulting situation
RTG	Research Task Group
SIEM	Security Information and Event Monitoring
SSMs	state space models
Sti	stimuli
STIX	Structured Threat Information eXpression
TAP	Technical Activity Proposal
TAXII	Trusted Automated Exchange of Indicator Information
TCP	Transmission Control Protocol
TTPs	techniques, tactics, and processes
TTX	table-top exercise

UDP	User Datagram Protocol
USB	universal serial bus

1 (PDF)	DEFENSE TECHNICAL INFORMATION CTR DTIC OCA
1 (PDF)	DEVCOM ARL FCDD RLD DCI TECH LIB
1 (PDF)	DEVCOM ARL FCDD RLD FT T BRAUN
1 (PDF)	US NAV ACAD D MCCALLAM