Threat Genomics

An evolution and recombination of best-available models and techniques for characterizing and understanding computer network threats.

Jonathan (Jon) Espenschied, Network Security Advanced Analytics
Angela Gunn, Security Response Communications
Microsoft Trustworthy Computing (TwC)

Abstract

Many threat models seek to determine or label the state of an online attack from discrete events. Static threat models based on measuring and correlating these discrete events are useful for detection and identification of like events and clusters of events. However, the responsive and predictive utility of static models is limited by the lack of timing and behavioral data. We seek to determine the nature of an attack from recognizable sequences of states and transitions between states, and use that data to address practical questions such as “What’s going to happen next in an ongoing attack?” or “What is the likely type of targeted asset given a sample of behavior?”

In previous work we define “threat sequences” that combine organizing concepts such as “relative superiority graphs” and the “kill chain” with quantitative and qualitative metrics such as threat timing and targeting. Building on this, we assemble a model for threat activity that allows security events to be organized into normalized base categories of activity, and propose extended metrics for transitions between those states or categories. By linking state transitions into a library of common sequences, we can apply sequence analysis techniques and parallel processing refined by biological sciences, resulting in a practical ability to anticipate unseen events and patterns based on sufficient fragments of a threat sequence.

Further, we show the efficacy of mapping activity and time data to established behavioral metrics and existing large collections of data for social behavior in order to derive information about intruders and operators of large-scale sophisticated attacks. This in turn results in improved ability to differentiate between concurrent activities and to more accurately identify adversaries. Practical application of these tools and metrics to risk reduction, implementation in current security information management software, and further research are discussed.

This document is intended to provide a framework and promote discussion about threat metrics and information interchange. It contains descriptions of events and entities drawn from public sources, which are provided "as-is" for demonstrative purposes only. Information and views expressed may contain errors or change without notice. You bear the risk of using it. Microsoft makes no warranties, express or implied, with respect to the information provided here. This document does not provide you with any legal rights to any intellectual property in any Microsoft product.

Contents

Introduction and Foundations

The genesis of this model in its current “Threat Sequence” or “Genome” form has been slow and iterative. There is no revolution in this structure, and the authors and contributors expect (and hope) that it will continue to improve through use and discovery of areas where work is needed. Key ideas that formed the foundation of the Threat Sequence include both well-known models from technical (“cyber”) and military theorists, as well as empirical observations about what worked (and didn’t) when applied to real-world computer incident response.

Focus on Actors not Tools

Foremost is the focus on actors and not tools. The age when a sophisticated adversary could be observed or characterized using a single signature or detection is long gone. An adversary that repeatedly uses the same tool or exploit within a single analysis domain can be described as *not very good*. Competent and serious adversaries do not make the work so easy; instead we must abstract our technical signatures and alerts with a layer of labels for actions and ask “What is the adversary doing?” Or “What are they after?” Only then do we have the right mindset and toolset to find an adversary that never retraces steps or reuses detected tools.

Influences from Relative Superiority and the Cyber Kill Chain

The Relative Superiority model developed by Colonel William McRaven provided a set of key ideas for graphically modeling the activities of sophisticated military operations over time. McRaven’s model showed the probability of mission success (Y axis) over time (X axis) from the perspective of an attacker. Over a series of analyses this work demonstrated that several consistent principles were common to successful sophisticated attacks, including knowledge of the target, and the simplicity and speed of the initial steps.

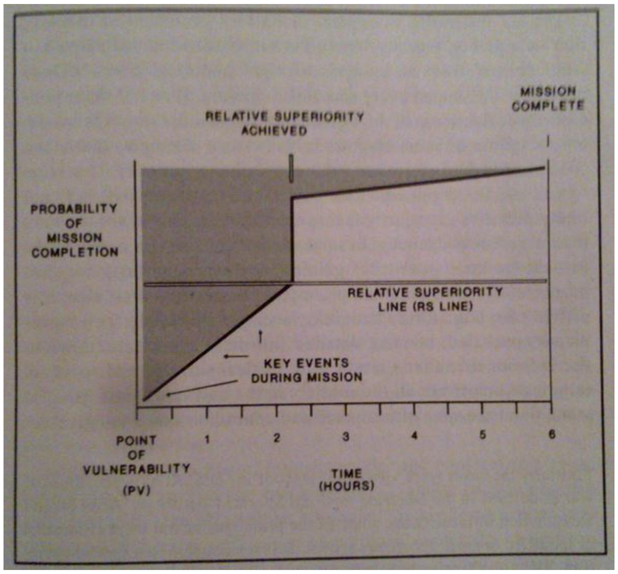


Figure 1. An example of McRaven’s Relative Superiority graphing

While the Relative Superiority model's metric for probability of success, defined as the point at which relative superiority >50%, maps generally to information- and online-centric attacks (i.e. an attack is more or less likely to succeed), the specific percentage-based metric does not. Military history provides a rich library of methods and procedures to measure a given tactical situation, allowing quantitative assessment of events on a timeline.

However, models for online attacks are still relatively novel, and it's hard to defend assertions such as "compromise of a non-privileged account means 33% likelihood of attack success" or "maintaining ongoing access corresponds to 85% tactical superiority." McRaven's concept of operational superiority holds, but quantitative metrics simply don't have hard meaning in this context. In its place we pursue a sufficiently accurate interval metric or consistent set of categories that show a series of actions that often have a dependent relationship or can be arranged in a rough order of time.

The classic notion of the military "Kill Chain" includes four stages of an attack and the dependencies between them: (1) Target identification, (2) Force dispatch to target, (3) Decision and order to attack, and (4) Destruction of target. In a series of well-known postings by Mike Cloppert and a related paper, this military model was extended to show stages of computer network attacks and the dependencies between them. The result was the "Cyber Kill Chain": (1) Reconnaissance, (2) Weaponization, (3) Delivery, (4) Compromise/Exploitation, (5) Command & Control, and (6) Exfiltration.

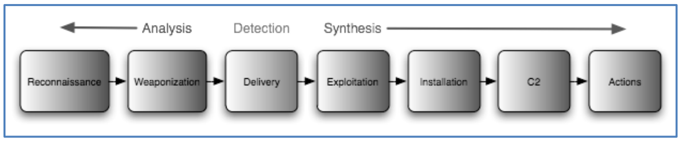


Figure 2. An example from the Cyber Kill Chain model and analysis

This was a large step forward for characterizing attacks, especially for activities that fit the military concept of a maneuver: an actor uses a distinct capability to perform a discrete action on a target entity or asset. However, persistence, lateral movement, concealment, and some other activities are hard to characterize and pose some difficulty within the model.

Addressing Persistence, Movement, Transitions and Graph Theory

Another set of models, such as one proposed by Trend Micro for characterizing targeted attacks on "cloud" infrastructure, handle activities that are difficult to represent in the Cyber Kill Chain. For example, a labeling model was proposed with: (1) Reconnaissance/Targeting, (2) Delivery Mechanism, (3) Compromise/Exploit, (4) Command & Control, (5) Persistence/Lateral Movement, and (6) Data Exfiltration. However, targeting and acquisition are still not addressed, making some activities difficult to correlate with similar actions. To handle these, we started to combine and extend similar concepts to handle real-world events from our records. The resulting set of "base" types of action are the minimum necessary to accurately characterize an attack for our purposes.

This set of 10 possible and often-sequential components of an attack together with their timeline form a "Threat Sequence." This set of possible activities can be used to describe most any whole or partial computer network attack.

The components are defined as “Base Types” of the sequence:

1. Reconnaissance
2. Commencement
3. Entry
4. Foothold
5. Lateral Movement
6. Acquire Control
7. Acquire Target
8. Implement/Execute
9. Conceal & Maintain
10. Withdraw

From Labels and Patterns to Sequences

The ten base type actions are not always present in each attack, but they are common in the sense that any attack can be described – using these conceptual buckets or building blocks – with sufficient precision to allow meaningful analysis. In basic graphical examples, these base types of attacker’s actions can be shown on a vertical axis, with a timeline on the horizontal, usually with the point of discovery as T=0. Also key is backward-compatibility and the ability to cleanly integrate with other models such as the kill-chain and various detection frameworks.

Certain Microsoft efforts, such as Project Broad Street led by Adam Shostack, also led us to take a closer look at the syntax of the sequence of activities and the transitions between them. When the nature of a disease and mode of transmission (the ‘tooling’ for a disease) are unknown, sometimes all we have to go on are the common actions and environment in a state transition. This led us to examine the transitions into and out of common stages or states of an attack, which in turn gave us insight into types of information to associate parallel or sequential activities that might otherwise appear unrelated.

There may also be “extended” action data that helps with remediation (e.g. details of a vulnerability and 0-day exploit), or information that is contextual and required for historical purposes (e.g. the political backdrop behind Stuxnet), but a Threat Sequence of base types of action tell the story with sufficient detail for our purposes.

From Sequences to Genome and Phenome

Iterative refinement and recombination led to a composite system that can be mapped to most others, including backward compatibility for those using the kill-chain model and its variants. Known earlier as the “attack curve,” this Threat Sequence model is intended to provide the best-present balance between precision and implementability, between extensibility versus backward compatibility, and between theoretical elegance and practical use.

If we have in hand a solid language for characterizing threats and an understanding of historical events and patterns, we have put a good and rich body of semantics into an extensible and durable syntax. What we get out of it, then, depends on how well we utilize it, and whether we can maintain a good balance between theoretical elegance and practicality.

The “genomics” analogy between threat pattern modeling and biological genetic pattern analytics appears to hold, allowing us both inspiration and direct borrowing from that realm’s techniques for detection and recognition. From genetic sequencing, we can look at incomplete data or fragments of the Threat Sequence, and use objective data or “hard markers” to predict the missing pieces of the patterns. In many cases we are also in a position to observe the expressions of human actors' behavior – observable characteristics that form a parallel "phenome" to the hard markers. Those behavioral "soft markers" can help to predict missing pieces of the same attack patterns, improve our confidence in pattern matching, and differentiate between groups of actors with similar tooling and goals.

I. Base Types of Action

The purpose of modeling an adversary's actions is to get as complete a picture as practical, so that it can be understood and used in the future. Specifically, when considering something composed of many parts, we want to coherently characterize related components and their relationships. For computer network attacks, the components are "base types of action" towards a functional goal, and attributes associated with the activities.

Forest and Trees

The threat landscape has changed radically in the past decade. Network attacks of the late 1990s or early 2000s were often technical attacks that could be countered with technical defenses. No more; modern attacks are more likely to target human operators, relying on social engineering or deception to compromise credentials, then proceed with other stages of an attack using valid credentials to access authorized resources. We have to be able to detect and respond to changes in behavior, yet many preventive and detective defenses are still oriented toward solely-technical attacks, lost in malware signatures and tool analysis, and ignoring detectable behavioral changes.

Our functional goal is to refine a tagging system or taxonomy that characterizes patterns sufficiently for analysis and comparison with past and present actions. However, before attempting to gather objective attributes and build a sufficiently accurate characterization model, it is important to consider what not to do.

Avoiding Subjective Attributes and Judgment

Much of the current body of work on modeling of advanced persistent threat (APT) or determined and sophisticated adversaries seems to either take the form of overextended technical analysis of tools, or originate from military human intelligence (HUMINT) applied to technical artifacts. In some cases, analysts appear to be convinced they have found a distinct human adversary based on a technical signature that may be common across disparate groups of actors. In other cases investigators may conclude that multiple attacks are tightly related based on finding common code in tools, despite evidence that code reuse is prevalent even between nation-state actors engaged in active conflict. Anthropomorphism, extrapolation, and wishful thinking are rampant.

Sufficient Characterization with Objective Attributes

In contrast, much of what we present here is an attempt to follow the common traits of signals intelligence (SIGINT), which when managed properly permits only the observable into analyses. There is little or no trust of informers, few manual sources, and little speculation about code ninjas or black helicopters. To pursue characterization and completeness of a series of actions, we prefer to only look at similar historical actions, which are often an excellent indicator of future events and sometimes the only rational guidance one can find. In every case, we try to stay within the limits of reasonably objective metrics.

Ten Base Types of Action

In development and analysis, ten components provided sufficient precision for detection, analysis, and other purposes -- without requiring collection by analysts of subjective information such as intent, or encouraging quantitative metrics with exaggerated precision. These ten components of a sequence are defined as “Base Types of Action”:

1. Reconnaissance
2. Commencement
3. Entry
4. Foothold
5. Lateral Movement
6. Acquire Control
7. Acquire Target
8. Implement/Execute
9. Conceal & Maintain
10. Withdraw

As noted earlier, these types are also backward-compatible with other models such as the kill-chain, and are designed to cleanly integrate and various detection frameworks. For example, the first four stages of the “Cyber Kill Chain” model map one-to-one with Threat Sequence base actions, and the latter six Threat Sequence base actions cover the remaining kill chain data with increased specificity.

Note that while there is an apparent natural order based on conceptual dependency, they are numbered here for reference only. In practice it is rare for a real attack to include all base types sequentially, and improbable in the extreme to detect an attack at its initial reconnaissance.

1. Reconnaissance

Reconnaissance includes knowledge-gathering activities such as area target identification, point target identification, target refinement to identify assets and entities, technical mapping of an environment, probing for vulnerabilities, and any other activities that are observational and can be applied to future adversarial actions. Note that this type of activity is rarely the first detection of an attack, so it may be useful to implement timestamps in an event database that are relative (Δ-T, in retrospect for this activity) to the first detection.

2. Commencement

Commencement is the point at which an adversary deploys tools or actions. The start of an attack does not necessarily mark the end of (1) Reconnaissance, as observation and knowledge gathering may continue, but the adversary has passed a threshold of observation and identification of vulnerabilities, and proceeded into engagement. Indicators in this stage may include discovery of intrusive scanning, active probing for vulnerabilities, phishing campaigns, or weaponization (per the Kill Chain model) of data or technology gathered earlier.

3. Entry

The traversal of a primary security boundary or network border can be characterized as entry by the entity responsible for that border. This traversal may take many forms, including a first successful entry or direct breach of a system, an 0day exploit against an external-facing application, a social engineering compromise of a person with basic privileges, or physical intrusion into controlled but not critical areas. Broadly interpreted, this includes any type of action that results in an adversary having basic access to one or more internal environments. The kill-chain “delivery” stage often maps to this type of action.

4. Foothold

The most common form of foothold is an adversary's breach of a system with low privileges or local credentials. The adversary is able to stay resident in the environment (actions that give only access would be characterized as multiple entries), but control or credentials used to stay resident are not privileged beyond the compromised system or applications. In the kill chain, this may be characterized as compromise/exploitation involving non-privileged hosts or credentials. Indicators may include localized compromises of a system, compromise of employee credentials, uploading tools, installation of remote access tools, and attempts at privilege escalation.

5. Lateral Movement

Lateral movement is the freedom to explore from within. When an adversary has the ability to move beyond a system or non-privileged environment in which they've established a foothold, the immediate network and adjacent logical environments are at risk. Indicators of this type of activity include detection of movement in network flow data, proximate reconnaissance activities within the environment, compromise of adjacent non-privileged or non-key systems and networks, and attempts to "pivot" controls or credentials to areas or assets of higher value. In McRaven's Relative Superiority model, this corresponds to the lower threshold of control superiority in an environment; the adversary has freedom of movement but has not yet established area control or crossed the "RS50" mark. Extended periods of activity in lateral movement with little progress may indicate insufficient reconnaissance, ineffective capability, or poor targeting depending on other factors.

6. Acquire Control

Acquiring or establishing privileged control in an environment allows an adversary free movement and access to assets and resources. It may be possible to detect and observe an adversary taking these types of actions, but it becomes difficult to dislodge an adversary that wields privileged access and control over a whole environment. This corresponds to the kill chain command and control stage, and to the upper side of McRaven's Relative Superiority (RS50) threshold. With an adversary in this stage, their attack is more likely to reach a state of success than not.

Indicators that an adversary has established area control may include loss of privileged administrator credentials, elevation of privilege in other accounts, and installation of backdoors or remote access tools on systems with sensitive data or applications. With certain classes of attacker, this type of activity may not mark the end of (5) Lateral Movement; in fact lateral movement may markedly increase for those intent on gathering all available assets or when first-line attackers are done and pass the access (or sell it) to lower-level actors.

7. Acquire Target

When an adversary can access a target asset, neutralize point target protections, and otherwise consolidate control over an asset, resources, or capabilities, that adversary is said to have acquired a target. It may be useful to compare (3) Entry and (4) Foothold (control of a non-privileged system and environment respectively) to (6) Acquire Control and (7) Acquire Target (control of area and point targets respectively). Indicators that an adversary has control of a target may range from deactivation of key administrative controls, to filtering and compression of data files for future transport, or PKI system compromise.

8. Implement/Execute

Implementation of a process, or execution of attack code on an acquired target, mark this base type of action. Indicators may include alteration in the function of key applications, consolidation and integration of control, exfiltration or destruction of data, or even communication of demands for ransom or other actions. An adversary's operations on a target generally correspond to the Kill Chain model's stage of exfiltration or actions on objectives.

9. Conceal & Maintain

With sufficient control of an environment and its contained systems, an adversary may be able to remove or alter security logs, implement decoy(s), periodically check backdoor access, or otherwise conceal and maintain their presence and activities. Indicators of this type of activity are often indirect; they may include corrupt or missing logs, aberrant behavior of systems or applications, or loss of access by administrators. In worse cases, indicators may come from business processes, financial checks and balances, or technical controls outside the organization. At this stage the target environment may be described as chaotic, as the owner has lost or has limited awareness of activities and events, leaving security responders to take action on symptoms rather than causes.

10. Withdraw

Withdrawal shows an attacker or adversary has completed its significant actions and intentionally departed *under its own terms or capacity*. Note that an adversary driven out or removed is not said to have withdrawn; in such cases the actions toward removal are taken by the environment's owner and not the adversary. Indicators may range from validated logs indicating conclusion of actions, removal of previously detected tools and access, to key indicators from external parties. In more serious cases, the withdrawal may be indicated by even more disruptive actions such as self-destruction of remote access or concealment tools, or wholesale destruction of data on affected systems.

II. Structure & Sequences

Predicting the future by extrapolating the possible is a path down which lies madness. Those seeking to avoid paranoia can find solace in history: A well-formed database of historical events is useful for estimating present or subsequent moves of an adversary based on patterns of past actions. While sophisticated adversaries may use a given tool only once or produce a distinctive 0day when necessary (or frustrated), the behavior of operators moving through a sequence of base actions often remains consistent for long periods and over multiple campaigns – consistent enough for our purposes to model, correlate, and build mechanisms for improved detection and response.

This gives us a path forward, through a small list of tasks:

- Find or build a collection of relevant events, and build a library of their threat sequences.
- Assess and evaluate what controls are present in the environment, and build a matrix to show which can give indication of activity or alerts in each base type of action. (In several cases, such as lateral movement or acquisition of targets, there may be no purpose-built control and a set of compensating controls and correlated analysis may be the only option.)
- Evaluate the common attributes available in each type of alert, to improve or enable the process of correlating activities.

Building a Library of Known Threats

To measure and understand new events, it is necessary to have a knowledge base from which to start. Even organizations that have not experienced a major attack can create a small analytical library of events. With logically sufficient building blocks in hand, it is possible to look back to historical events and define patterns that can help with detection, understanding, and response to new events.

Finding Relevant Events

The initial library of events should be relevant to the group or organization; this can be by assets at risk, business type, size, financial attributes, supply chain, or any other major factors that relate to actual risk to the organization. For example, a bank might select well-documented network intrusions into credit card processors, but not nation-state attacks against government agencies. On the other hand, a regional energy utility might be concerned with both credit card processing because of its payment systems, as well as nation-state attacks that are relevant to operating critical infrastructure such as hydropower and flood control systems or a nuclear power plant. A variety of creative activities may be employed to gather relevant incident data; we employed games, trading cards, and other exercises to positive ends.

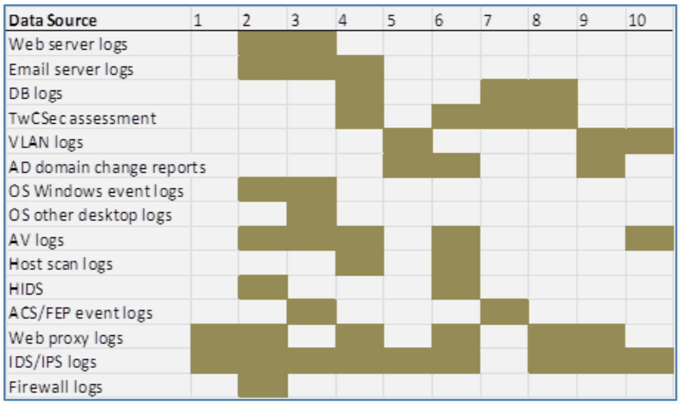
Relevant events can be cataloged, and the pattern of actions for the event documented using the base types of action. Where possible, the timeline of events has proven useful, as described in later sections. Not every type of action will be present in each relevant event, and it may be common to see jumps, omissions, loops, or other non-linear progression.

Linking Base Actions by Attribute or Time

Base types of action all have some comparable attributes, such as the mode of detection (often an "indicator of compromise" or similar); time of detection and duration; number of concurrent, sequential, or periodic actions; and a source, destination, and vector (which often shows an immediate but not end target). The data used to associate detected actions depends on the environment, but will usually include all of these basic attributes. Extended information may be quite detailed, ranging from packet captures and network flow data to threat intelligence data.

Data Sources and Indicators

For network attacks in the present and future, a variety of alerts and indicators may be available. Some will be direct and trivial to interpret; when an anti-virus or host intrusion-detection system sends an alert, there is usually a problem on a computer system that can be characterized as a foothold, or acquiring control of a higher-value system. Other activities, such as lateral movement and concealment, may require combinations of alerts, careful configuration of heartbeat signals, or correlation engines to ascertain their occurrence.



Data Source	1	2	3	4	5	6	7	8	9	10
Web server logs		■	■							
Email server logs		■	■	■						
DB logs				■			■	■		
TwCSec assessment				■		■	■	■		
VLAN logs					■				■	■
AD domain change reports					■	■			■	
OS Windows event logs		■	■							
OS other desktop logs			■							
AV logs		■	■	■		■				■
Host scan logs				■		■				
HIDS		■				■				
ACS/FEP event logs			■				■			
Web proxy logs	■	■		■		■		■	■	
IDS/IPS logs	■	■	■	■	■	■		■	■	■
Firewall logs		■								

Figure 3. An example matrix of data sources for categories across the Threat Sequence

The combination of data sources or "feeds" are different for every organization, but a common approach is to make a simple matrix of relevant controls compared to the base types of action. In its simplest form, this can identify areas in which there are insufficient controls, before any detailed threat modeling is undertaken. Subsequent iterations can be used to refine the controls for detecting each type of action. Tabletop exercises based on the library of known threats can be used to validate that appropriate controls are in place.

NOTE: Certain further developments in this area may be available under NDA or similar agreement.

Common Attributes to Relate Base Actions

Constructing Threat Sequences requires some way to associate events. In many cases the most expedient option is to use a commercial correlation tool, which provides a consistent format for intake of event data so that the events may be processed and connected to other events or states. However, this is certainly possible for those without such tools or wanting to pursue detailed correlation through custom tools or manual processes. Finding and defining relevant attributes for each base type allow us to connect, group, and ultimately correlate activities that might otherwise appear unrelated. These attributes will vary for every organization and environment; those in the following sections were useful in our inquiries.

Commonly Gathered Attributes

We found the following starting set of attributes were useful as a baseline for most scenarios in which the threat sequence modeling was used to organize threat data.

- An identifier (ID) and optional name for automatable reference to the event or action
- Time detected, usually a marker of first detection set by an IOC
- Duration start, Y/M/D/H (or Δ- D/H time in retrospect)if different from time detected
- Duration end, or last known/confident detection
- Base type of action, usually estimated by analyst or normalization rules
- Source of alert or detection, specific or in aggregate with ID that allows traceback
- Targeting, including evidence of randomness or selection by opportunity, area, sequence, or point
- Estimation of operational and technical sophistication
- Indicator of Compromise (IOC) record, if available
- IDs of all involved source/destinations, whether system, account, or application
- Vector, showing incoming, outgoing, stasis, or lateral movement; avoiding intermediate guesses of victimhood or attribution

Extended Attributes

The following information is commonly available, but is often focused on a specific vulnerability or incident. When adjusted for the environment, these pieces of data may provide additional insight and ability to correlate with other actions or events.

- Time in relation to potentially related Base Actions
- Evidence of human behavior, including parallel or sequential actions, decisions, escalation, coordination, defacement or other markers, and other behavioral attributes
- IOC or other alert record
- Alert source and type
- IPv4/6 and any DNS records for involved entities
- IP flow or trace data, or other captured data in the alert
- Target asset sensitivity or entity access level; a suggested basic nomenclature is:
  - Low: Public or low business impact data for which integrity outweighs confidentiality
    Higher range: Negotiable assets (money/financial assets which may be insured)
  - Medium: Confidential or medium business impact data
    Higher range: Tools, code, credentials, or data which allows elevation
  - High: High business impact data, such as critical trade secrets and classified data
    Higher range: Assets affecting human life and safety, or compartmentalized information
- Organization type, usually by industry, size, or business relationship; such as:
  - General populace/individuals

- Education, research, and other independent nonprofits
- Technology and telecom organizations including software, hardware, integrators, and operators
- Industries including service, retail, manufacturing, and materials producers
- Infrastructure and transport including all utilities
- Finance including banks, CU, credit, transaction processors and financial NGOs
- Government including all fed/state/local civilian agencies, domestic intelligence, and law enforcement
- Military including geopolitical actors, international intelligence and some NGOs

Preparing for Analysis

With only a basic set of events characterized by base action and time, it is often possible to see notable patterns even at the outset of analysis. For example, persistence without progress in the first few stages of an attack (i.e. compromise of low-privilege credentials or foothold systems without the immediate ability to elevate and proceed to other activities) is a reasonable indicator that the adversary did not conduct extensive reconnaissance or detailed targeting. Conversely, fast attacks that are wildly successful are often revealed to have long and detailed reconnaissance upon investigation.

Other common findings include higher operational sophistication in line with higher value of assets, but no particular correlation between technical sophistication and targeting effectiveness or persistence. (Sophisticated adversaries are just as often content to use mundane tools carefully; the use of 0days is not a good indicator of operational sophistication. Broadly speaking, thieves and vandals don't behave like nation-state actors or industrial spies. Simply mapping relevant historical events against the base types of action may prove useful in context.

In addition, it is important to recognize common fallacies about modern network attacks. Similar-appearing attacks may be very different in technology and management. Not every attack follows the same dependencies, and some attacks may skip entire types of action. By limiting data input to observed actions and patterns, and not the possible future or estimated capabilities of an adversary, more dependable analysis templates can be created. This sets the stage for analysis described in the following sections.

III. Genome

With sufficient information, sequences of activity from a common actor or types of adversaries can be detected from "hard markers" and recognized with reasonable confidence. If the adversary is known, actions can be linked and a synopsis of an attack constructed from the attacker's base actions over time.

If the attacker is not obvious or the sequence is incomplete, detected adversarial actions can be categorized, and then associated by common attributes of each action or emergent patterns. Whether or not that association is sufficient to provide attribution or just differentiation from other potential actors, whole or partial sequences can be compared to known historical patterns. Those known patterns often provide useful indicators of where to look for undetected past or future activity.

The analysis of these Threat Sequences is similar in basic ways to genome analysis. Manual "sequence walking" and a variety of automated sequence mapping techniques (gratefully borrowed from genetic researchers with harder tasks than ours) may be used to sequentially compare an unfolding event to a series of fully-defined historical patterns, or even to perform such processes in parallel to address great volumes of data.

The genome analogy appears to hold through various methods of estimating or projecting the unseen parts of a sequence based on the observable fragments. With a library of threat sequences assembled, significant patterns can be isolated, adjacent or isolated ones identified, and when an alert triggers for a pattern, the detection thresholds for related or "chained" ones may be lowered for a period of time.

Constructing and Using Threat Sequences

What indicates that an intruder's activity is lateral movement versus compromise of another machine, or what differentiates compromise of network devices from concealment? What can we do with a series of well-defined threat sequences? Can the results be visualized or presented in a way that improves the speed and accuracy of human understanding? The test of any tool's improvement is how well it performs in real-world conditions.

Setting Up and Using Detection Metrics

Indicators or inputs have to be defined for each base type of action. Most host intrusion detection systems, for example, have rules and alert thresholds for events on a host, and when those criteria are satisfied they send an alert that could signify Foothold or Acquire Target depending on the nature of the host. Likewise the logs from border firewalls and analytics from web applications may provide a baseline from which Reconnaissance may be derived, and internal security log analytics servers may be used to spot sensors or logs that have been tampered with, indicating Conceal & Maintain actions.

The matrix of available controls is different in every enterprise. To set up controls and metrics for detection, the environment owners should take an inventory of the available controls, and review their applicability to each base type of action. For the purpose of characterizing an adversary, focus should remain on the output of detective controls, or the status of controls intended to deny an adversary's actions. (Remediation and self-help such as items in the the DoD action matrix are outside our scope.)

The desired result is a mapping or matrix of one or more detective or preventive controls for each base type of action. This mapping may be refined by identifying specific alerts, combinations of alerts from multiple sources, or correlated events that differentiate between actions. For example, Acquire Control might be indicated by a NIDS alert about moderate anomalies on an internal network combined with a series of HIDS alerts on specific machines, but the same NIDS alert combined with directory logs showing unusual connection attempts might indicate Lateral Movement.

To move from malware detection to behavioral recognition, we have to establish what constitutes “normal” over a sufficiently long period of time to account for normal variations in the rhythm of business. Only then can we begin to recognize the behaviors that don’t belong, and more specifically, the behaviors that would seem normal if only they were being performed in another context, against other assets, or in another pattern of time or movement.

Finding Recognizable Patterns in Common Sequences

With a good baseline in hand, data from past intrusions or other security events can be used to build a library of known Threat Sequences. As noted earlier, it may be useful to consider a Threat Sequence as a signals-driven construct, in contrast to the intelligence-driven process of the Kill Chain whereby actors and a chain of dependencies are identified: It is perfectly acceptable to dump all known activity from all known sources into a database, and to correlate common and extended attributes to find emergent patterns.

The dataset may be too large to directly detect Reconnaissance activities through deep analysis of public web server logs for a large enterprise, and following signatures for the use of specific attack tools or malware is often useless in tracking a skilled adversary. However, if the logs and alerts are correlated by base type of action, one can recognize more serious events by adjacent activity in a Threat Sequence. For example, anomalous account use may be understood differently if it is connected to a preceding malware alert on a machine associated with that account (a foothold or lateral movement) and a subsequent lack of normal update behavior (acquire control or conceal) regardless of the specific signatures for those events.

Visualizing and Understanding Threat Sequences

Nice graphics are not required for a deep understanding of data. However, in any situation where there are stressed humans subject to adversarial action and pressed to make decisions, any tools that aid understanding and response are of great utility. We aim to improve the detection and understanding of attack patterns so that the appropriate course of action can be selected faster and more accurately.

A recurring example specific to large knowledge-driven enterprises is the ability to differentiate between a less-serious incident with a smash-and-grab goal of monetary theft, and the more-serious intruder intent on doing significant damage or appropriating high-value assets, and remaining resident. One technique for visualization is presented in the earlier paper “Threat Behavior: Attackers & Patterns,” which shows the divergent behavior of thieves and vandals, versus the behavior of industrial spies and nation-state adversaries. (In that paper and the middle example below, the timeline for Entry, Foothold, and Lateral Movement is often much longer for the former types of adversary than the latter.)

Three approaches to usage and visualization are shown:

- Simple sequencing to characterize a maneuver or single attack
- Differentiating between complex attacks with threat sequences over time
- Advanced sequencing to recognize campaigns or multiple actors

Simple Sequencing to Characterize a Maneuver or Single Attack

For truly simple intrusions, defining common patterns is left as an exercise for the reader. The Common Attack Pattern Enumeration and Classification model is excellent for characterizing an adversary that has only one tool; where the attack is basically synonymous with a malware definition, it can help identify and quickly respond to repeats, copycats, and variants. An instance of the Cyber Kill Chain may also be used to characterize attacks with simple toolset and linear dependencies to an objective; its focus on intelligence is suitable for deep analysis of actions on vulnerabilities and network defense. In common military parlance, these intrusions can be thought of as comparable to the formal definition of "maneuver."

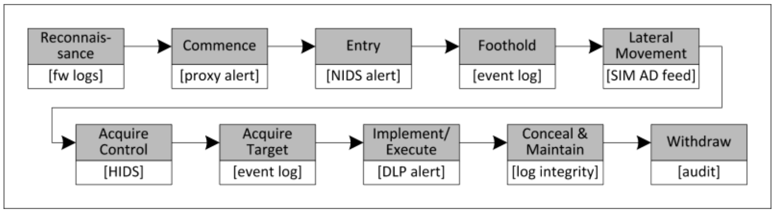


Figure 4. A simple sequence of actions in a security incident.

Representing a simple attack or maneuver with the Threat Sequence model's extended base types of action, together with the sources of alerts and other notifications, may be useful to get a basic handle on what happened during an incident. In cases where a control may have failed or not provided adequate information to inform responders about what was transpiring, this can be used to support remediation and longer-term technical mitigation.

Differentiating Between Complex Attacks with Threat Sequences

The patterns of activity for more sophisticated actors can often be recognized by commonality of targets, but the Threat Sequence model's introduction of time as a pervasive attribute, and a more-detailed set of base types of action, provides an easier method for recognizing and differentiating between major types of attackers. This "attack curve" form of the Threat Sequence model shown in Figure 5 can be manually generated from limited data, and still provide substantive utility in recognition.

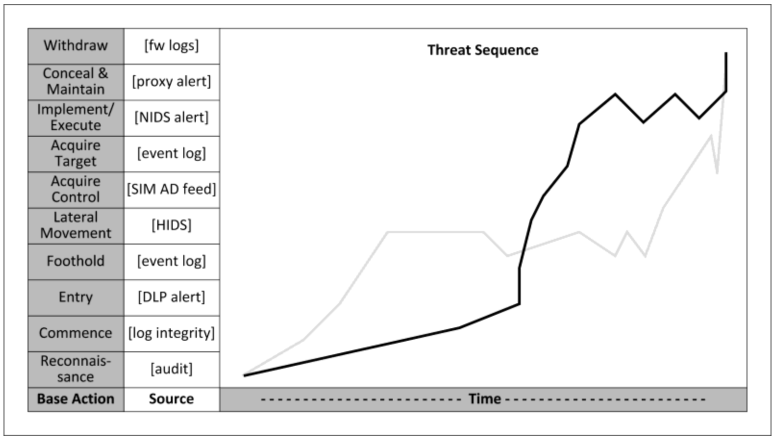


Figure 5. An “attack curve” visualization of a Threat Sequence.

In Figure 5, the Threat Sequence of a sophisticated adversary’s actions are mapped over time. In the background, a pattern showing the aggregate typical behavior of vandals or credit card thieves is shown. The contrast makes it clear that the foreground example is different; it fits the typical behavior for industrial spies and nation-state actors, which are often a serious threat to high-value information and basic control over an environment. By introducing time and making it an integral part of the Threat Sequence, similar alerts over different periods can indicate very different types of adversaries, and often one can often differentiate between them within the first few base actions of the Threat Sequence.

Advanced Sequencing to Recognize Campaigns or Multiple Actors

The limitations of source data are significant and often disappointing. For example, with present technology it is much harder to qualify certain alerts as signifying Lateral Movement of an adversary, versus a properly authorized privileged user who happens to be doing something new. In addition, sophisticated actors may have teams of people (which doesn’t make for pretty line graphs as in the previous example), and the cessation of certain forms of activity may be as informative as others’ presence. To address this, large-scale automation can be leveraged, with sophisticated methods for correlating sequence fragments and for visualization.

For example, in Figure 6, concurrent activities signifying Reconnaissance or Commencement of an attack may not be particularly interesting in a noisy environment, but if several stop all in parallel, it may signify the success of a lower-profile activity (likely ignored or filtered) in an adjacent or dependent category of action. By using known patterns or templates for Threat Sequences, significant actions below the detection threshold may be found by lowering alert thresholds for a specific time period and focusing on data sources applicable to missing fragments based on historical or known Threat Sequence patterns. Such queries might otherwise be computationally infeasible, or swamp security administrators with useless data.

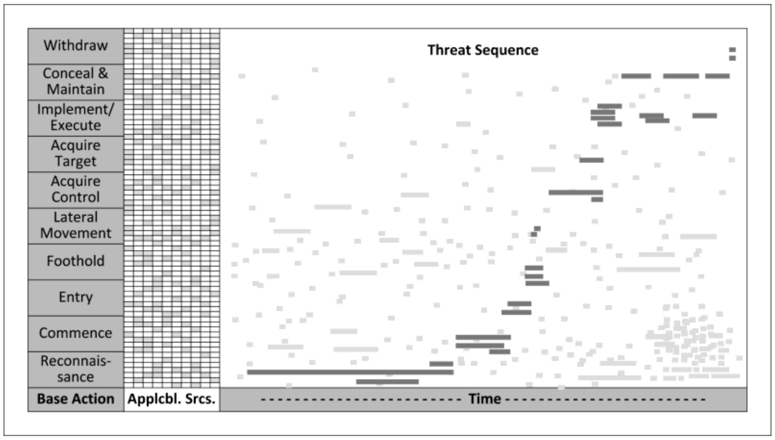


Figure 6. A source-mapped Threat Sequence against a background of other detected activity.

In addition, this sort of correlation and visualization may make it easier to differentiate between multiple actors by connecting related and dependent actions, such that sophisticated but subtle adversaries may be identified even against a noisy background. For example, the lower right of Figure 6 shows a cluster of alerts typical of a large network on the receiving end of Anonymous' ire, or that of any other large cluster of individual actors. This might, and often does, distract analysts from a smaller but more advanced threat. While such big adversaries may be powerful and destructive, an immense but lightly structured crowd is fundamentally different in behavior from small groups of financially motivated attackers, and both are distinctive in comparison to industrial or nation-state actors with specific targets, project managers, and salaried staff.

By embracing "big data" with a hybrid of quantitative data and a sufficiently advanced qualitative approach to finding patterns, we can use hard data from the past to recognize present and likely future events. We can improve our understanding of the data available, and in turn ensure that our security controls align and provide good coverage for the types of activity that cause actual risk.

IV. Phenome

Objective data or “hard markers” that indicate the state or sequence of an adversary’s actions may in most cases be enough to inform us about patterns or fragments of the adversary’s behavior, and match it to known patterns. However, in many cases we are also in a position to observe the expressions of human actors’ behavior – observable characteristics that form a parallel “phenome” to the hard markers. There is frequently useful information about apparent human behavior in the transition between each base type of action, such as the number, repetition, parallelism, and persistence. Those behavioral “soft markers” can help to predict missing pieces of the same attack patterns, improve our confidence in pattern matching, and differentiate between groups of actors with similar tooling and goals.

Origins and Goals

Phenome

Although the term causes consternation among some genetic researchers, “phenotype” is the result of an entity’s interaction and state within its environment, resulting in observable characteristics. Often attributed to Francis Galton (a cousin of Charles Darwin) who opened a “Anthropometric Laboratory” in the early 1880’s, this approach recognized that genes were not deterministic in the real world. Environmental influences, actions of others, disease, vaccinations, immediate history, and other factors tended to obscure direct measurement of genetic attributes by observation. Complicating matters is that certain inherent traits may be ineffective or turn off in certain environments – a pattern that is directly analogous to malware and tools on related but not identical platforms.

Measurement of externally observable traits may sometimes be more effective and expedient than the extra time and effort required to quantify or even to qualify inherent traits. That an adversary’s tool contains code derived from Stuxnet may be only a curiosity, while the adversary’s repetitive behavior focusing on a particular difficult target may disclose critical information about subsequent activity.

Differentiation not Attribution

Although this model does not address specific attribution of attacks, these additional behavioral indicators can be used to refine differentiation between clusters of activity, and provides another method to find serious adversaries in otherwise-distracting background noise. Our focus on differentiation is a current end-goal; although the material later in this section uses examples of specific countries or groups, we do not attempt to attribute activities to named entities in this model . Attribution based on these soft markers is directly addressed in recent work from Char Sample, whose ideas and shared information were a major factor in this portion of the model.

Data Sources for Observable Expressions of Behavior

Two actors attempting the same attack, even with similar tools, goals and timeframe, may still differ in their approach due to cultural and organizational differences between the two. One may prefer a long period of reconnaissance followed by a quick intrusion and short period of concealment because the cost of detection is perceived as high, while the other may prefer a longer period of persistence in the end stage because attribution is perceived as low risk, or even desirable after the major activities are completed. These variations in observable expression may have a cultural basis, an organizational basis, or a combination of the two. These are still measurable, though some switching of gears is necessary.

Back to McRaven's Principles

In his case studies on sophisticated military operations, McRaven called out six observable principles that he found to be present in all significant and successful operations – simplicity, security, repetition, surprise, speed, and purpose. Many of these "Six Principles of Special Operations" are key indicators of attacks that reached RS50 quickly and subsequently succeeded, and often manifest in fairly specific, measurable ways. In addition, McRaven's Relative Superiority graph visualizations, which were the basis for much of our early work, are meant to show them.

The key concept is that these attributes are relevant to the actions of sophisticated adversaries. McRaven noted that when any one of them was substantively absent, the operation or maneuver invariably failed to some degree. Likewise, types of action and transitions between types of action that must always be present for success of an online attack are of particular interest to us, as they are more likely to be present in the activities of the more competent and determined adversary. If there is some pattern more common to serious attackers and ignored by less-sophisticated ones, we want to focus on it.

Observational Sources

Data sources for observable behavioral attributes include many of the same attributes McRaven looked for in special operations. In the actions of sophisticated adversaries we can observe simplicity of direct action, management, apparently preceding action (planning and/or duration), contingencies, use of surprise, speed, evidence of group activity, timing around holidays, evident creativity in problem-solving, and many more. However, McRaven again cites the necessity of certain attributes, but does not suggest a metric.

Geert Hofestede, then an IBM researcher, gathered extensive social survey data from IBM employees in the late 1960s, and spent a large amount of time in the early 1970s looking for emergent patterns in the data from more than 117,000 people across 40 countries. At the time, this was the largest such effort at social metrics. In subsequent years, Hofstede continued the work at IBM and elsewhere, expanding the cross-national database and combining it with the work of others to cover surveys from more than a million people across 93 countries.

From this work, six indicators have survived years of analysis as behavioral indicators of culture. That is to say, people consistently behave in certain ways when making decisions or evaluating situations, and they are measureable in ways that show consistency within cultures and dissimilarity between them. As our interest is in differentiation, these metrics are of strong interest to us. Hofstede notes that this is not the same thing as being able to identify an individual person and attribute their culture or nationality – such as identifying an American by his or her actions alone – and in fact he has objected to such efforts in the past. However, we're content to borrow just the metrics tested over decades to show statistically significant differentiation.

Social and Cultural Behavior Metrics

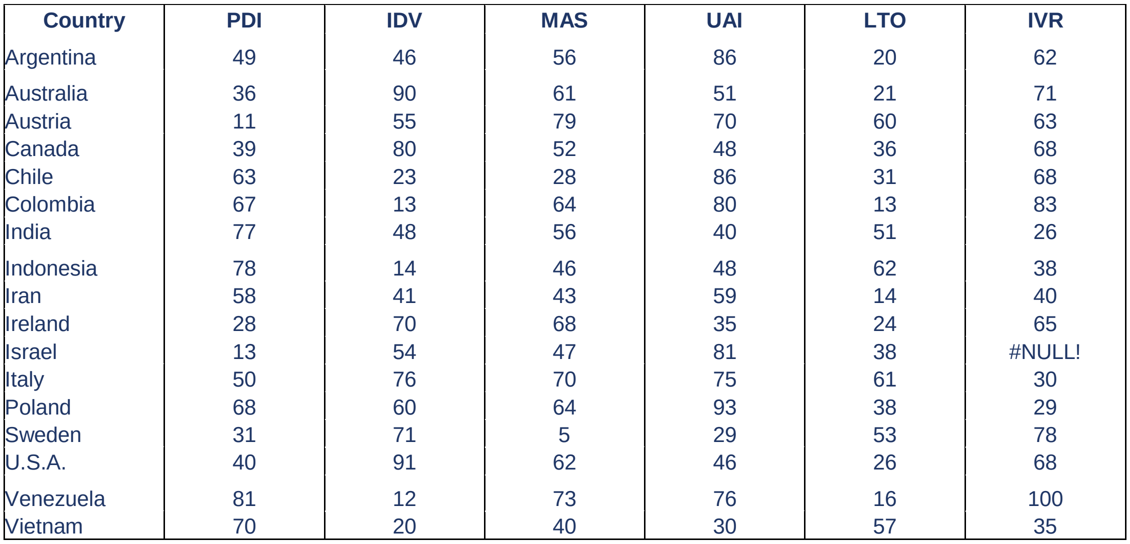
Cultural values may map poorly to individual actors, but the work from Hofstede, Bond, Minkov, and others show that the value metrics are useful as firm underpinnings and observable markers for how people act within groups, or in comparison to other groups.

Cultural Dimensions

The six cultural dimensions identified by Hofstede and affiliated researchers are:

- Power Distance Index (PDI)
- Uncertainty Avoidance Index (UAI)
- Individualism vs collectivism (IDV)
- Aggression (masculinity) (MAS)
- Long-term vs short-term orientation (LTO)
- Indulgence vs restraint (IVR)

Detailed descriptions of these indices can be found in Hofstede's "Cultural Dimensions" and related texts. In it he details the process by which the metrics were derived and gives exhaustive details and examples of the traits. The end result is a series of tables that show the "6 dimensions" metrics by country. A few of these are shown in the following example (including one country's score where there was not enough data):



Country	PDI	IDV	MAS	UAI	LTO	IVR
Argentina	49	46	56	86	20	62
Australia	36	90	61	51	21	71
Austria	11	55	79	70	60	63
Canada	39	80	52	48	36	68
Chile	63	23	28	86	31	68
Colombia	67	13	64	80	13	83
India	77	48	56	40	51	26
Indonesia	78	14	46	48	62	38
Iran	58	41	43	59	14	40
Ireland	28	70	68	35	24	65
Israel	13	54	47	81	38	#NULL!
Italy	50	76	70	75	61	30
Poland	68	60	64	93	38	29
Sweden	31	71	5	29	53	78
U.S.A.	40	91	62	46	26	68
Venezuela	81	12	73	76	16	100
Vietnam	70	20	40	30	57	35

Figure 7. Sample of the "6 Dimensions" data originating from Hofstede's research

Differentiation by Cultural Dimensions

It follows then if we can map consistently available and observable behaviors to a large and stable database that differentiates between groups, then we ought to map the data we have to the metrics that are available. For example, we map base types of action and transitions between bases in an incident, using qualitative behavioral metrics:

Power Distance Index (PDI)

Do parallel actors take the same actions? Are they using a playbook?

Uncertainty Avoidance Index (UAI)

Are attackers pragmatic? Do they adapt or keep trying failed attacks?

Individualism vs collectivism (IDV)

Is there an aversion to using "not invented here" tools? Tendency to follow group activity?

Aggression (masculinity) (MAS)

Is there direct reaction to being blocked or removed from a system? Are there markers for ownership or entitlement? Hostility toward remediation?

Long-term vs short-term orientation (LTO)

Is there an investment and intent to stay resident? Active maintenance or observation (not just time in a botnet)?

Indulgence vs restraint (IVR)

Is there defacement? Flair? A distinctive style or tendency to leave cryptic clues? Announcement of success, or petulance at failure?

A detailed matrix can be developed, with a set of specific or context-setting questions for the six dimensions, for each of the ten base types of action. While some questions are common across multiple types of action, this gives a minimum of roughly 50 qualitative questions – similar in nature to the surveys from which Hofstede drew his original data – from which to measure the behavior of adversaries engaged in an attack. Some examples of evidence for which questions can be drawn include:

- Site defacement is often an impulsive act attackers perform to assert their dominance over a network – until the defacement is taken down. It conforms to the base type Implement / Execute and indicates that the attacker has low Long-Term Orientation and a tendency toward Indulgence over Restraint.
- Some cultures are more aggressive (high-MAS) than others. When detected and thrown off the network by an administrator, some attackers may simply leave, while others may attempt to retaliate.

Our work has shown that a qualitative rating is useful in differentiation. Specifically, a 5-interval metric is used for rating each dimension, where the lowest and highest rating show a distinctive absence or overwhelming presence of an attribute, and a low/medium/high rating is available between the extremes. Ratings are measured relative to historical data, which means initial accuracy may not be strong, but improves over time.

A criticism of this approach has been to assert that all computer network attacks are inherently aggressive, with specific power relationships, and other common qualities, meaning that the corresponding metrics are useless. However, this has been shown not to be the case for most activities. In the few cases where an action is inherently matched to one of the dimensions, that single dimension can be excluded. This provides graceful degradation; By excluding inherent behavioral indices from bases where they are expectable, cultural artifacts begin to appear without the entire model falling down.

Differentiation by Organizational Dimensions

Another criticism of applying the Cultural Dimension data to the problem of online or even kinetic attacks is an assumption that all participants in such attacks are more homogeneous in their behavior as a participant in conflict than their behavior as influenced by culture and environment. While the question is valid, years of research indicate that the metrics persist and the dimensions are measurable even in extreme situations.

Nevertheless, research has been done to quantify observable characteristics that cut across cultures by way of specific organizations. As with cultural measures, some expressions are intrinsic to certain activities; other behaviors indicate specific value systems below the surface. Our pursuit of differentiation between adversaries then comes as close to attribution as we dare: Are attackers free actors? Are they corporate or military? Organizations often express their own corporate cultures in dimensions that can be measured relative to other organizations:

- Means vs goals (results orientation)
- Internally or externally driven (attitude toward customers)
- Easygoing vs strict work discipline (internal structuring)
- Local vs professional (identification)
- Open system vs closed (organization accessibility)
- Employee focus vs work focus (management philosophy)

These factors are the subject of ongoing research, and may provide still further refinement in the ability to differentiate groups of active sophisticated adversaries.

Conclusion

The problem of characterizing serious adversaries in computer network attacks – often referred to as Determined Human Adversaries (DHA) or Advanced Persistent Threat (APT) – is a difficult task, yet an essential one. Modeling their behavior, and tuning the model so that it is useful and accurate enough, is essential for earlier detection, faster response, and potential prevention.

The "Threat Sequence" model allows qualitative characterization and labeling of security events so that they may be normalized and correlated into a coherent whole – bridging the gap between individual signatures or detections to extended attacks involving multiple tools, targets, and modes of activity.

The 10 base types used to construct a threat sequence are labels or categories into which an attacker's actions can be sorted, with sufficient precision to distinctly characterize the attack for analysis, without forcing evaluation against quantitative metrics which may be unavailable or subjectively applied. Fragments of correlated activity can be compared against known whole sequences to potential subsequent actions, or to investigate otherwise-overlooked past activity.

This model for threat analysis stands on the shoulders of giants and contributing friends alike, combining long-standing and proven concepts with recent astute observations to move the problem space from daunting to approachable. This in turn enables practical applications of threat pattern data to deeper analytics and functional defense, and meaningful alignment and improvement of mitigating, detective, and preventive security controls.

Appendices

References

Scott D. Applegate, “The Principle of Maneuver in Cyber Operations”, George Mason University. http://gmu.academia.edu/ScottApplegate/Papers/1486050/The_Principle_of_Maneuver_in_Cyber_Operations

Mike Cloppert, "Attacking the Kill Chain," SANS CFIR, 19 Oct 2009. http://computer-forensics.sans.org/blog/2009/10/14/security-intelligence-attacking-the-kill-chain/

Jon Espenschied, “A Discussion of Threat Behavior: Attackers & Patterns” Microsoft Corporation and NATO CyCon, June 2012. http://www.microsoft.com/downloads/details.aspx?FamilyID= 8fbbe2a9-a548-4c69-a6d3-0b04a39574ea or http://www.ccdcoe.org/cycon/doc/20120501-Espenschied-ThreatPatterns-public.pdf

Matt Frazier, “Combat the APT by Sharing Indicators of Compromise,” Mandiant Corporation. https://blog.mandiant.com/archives/766

Geert Hofstede, “Culture’s Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations,” Thousand Oaks CA: Sage Publications, 2001 (second edition).

Geert Hofstede, Gert Jan Hofstede, Michael Minkov, “Cultures and Organizations: Software of the Mind,” New York: McGraw Hill, 2010.

Eric Hutchins, Michael Cloppert, Rohan Amin, Ph.D.; Lockheed Martin, “Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains” http://papers.rohanamin.com/wp-content/uploads/papers.rohanamin.com/ 2011/08/iciw2011.pdf

William J. Lynn III, "Defending a New Domain: The Pentagon's Cyberstrategy," Foreign Affairs magazine, Sept/Oct 2010. http://www.foreignaffairs.com/articles/66552/william-j-lynn-iii/defending-a-new-domain or http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA527707

William McRaven, “Spec Ops: Case Studies in Special Operations Warfare Theory & Practice,” New York: Presidio Press (Random House Publishing Group), 1995.

Mandiant Corporation, “M-Trends: The Advanced Persistent Threat” http://www.princeton.edu/ ~yctwo/files/readings/M-Trends.pdf

Mitre Corporation, “Common Attack Pattern Enumeration and Classification (CAPEC)” http://capec.mitre.org/about/documents.html

Char Sample, “Using Soft Markers in Attack Attribution,” presented at Shmoocon 2012. http://www.shmoocon.org/2012/presentations/Char_Sample_ShmooCon2012.pptx

Adam Shostack, “Security Breaches are Good for You,” presented at Shmoocon 2007. http://www.homeport.org/~adam/Security%20Breaches%20are%20good%20for%20you.pdf

Nart Villeneuve, “Trends in Targeted Attacks,” Trend Micro, 2011. http://www.trendmicro.com/ cloud-content/us/pdfs/about/wp_trends-in-targeted-attacks.pdf

Paul Wright, “Ten Stages of a Network Attack” (excerpt from "Oracle Forensics: Oracle Security Best Practices"). http://www.dba-oracle.com/forensics/t_forensics_network_attack.htm