TC 3-12.2.4.1

ANALYTIC SUPPORT TO DEFENSIVE CYBERSPACE OPERATIONS

JANUARY 2024

HEADQUARTERS, DEPARTMENT OF THE ARMY

Training Circular
No. 3-12.2.4.1

Headquarters
Department of the Army
Washington, DC, 17 January 2024

Analytic Support to Defensive Cyberspace Operations

Contents

Figures

Tables

Preface

TC 3-12.2.4.1 discusses the use of data analytics in support of defensive cyberspace operations.

The principal audience for this publication is analytic support personnel and data engineers who support defensive cyberspace operations. Commanders and staffs of Army headquarters serving as joint task force or multinational headquarters should also refer to applicable joint or multinational doctrine concerning the range of military operations and joint or multinational forces. Trainers and educators throughout the Army will also use this publication.

Commanders, staffs, and subordinates ensure that their decisions and actions comply with applicable United States, international, and in some cases host-nation laws and regulations. Commanders at all levels ensure that their Soldiers operate in accordance with the law of war and the rules of engagement (see FM 6-27). Commanders also adhere to the Army Ethic as described in ADP 6-22.

This publication uses joint terms where applicable. This publication is not the proponent for any Army terms. For other definitions shown in the text, the term is italicized, and the number of the proponent publication follows the definition. The mention of commercial products in this publication does not imply endorsement by either DOD or the United States Army.

This publication applies to the Active Army, Army National Guard/Army National Guard of the United States and United States Army Reserve unless otherwise stated.

The proponent for this publication is the United States Army Cyber Center of Excellence. The preparing agency is the Doctrine Branch, United States Army Cyber Center of Excellence. The technical review authority is the Cyber Protection Brigade, United States Army Cyber Command. Send comments and recommendations on a DA Form 2028 (*Recommended Changes to Publications and Blank Forms*) to Commander, United States Army Cyber Center of Excellence and Fort Eisenhower, ATTN: ATZH-OPD (TC 3-12.2.4.1), 419 B Street, Fort Eisenhower, GA 30905-5735, or e-mail to usarmy.eisenhower.cyber-coe.mbx.gord-fg-doctrine@army.mil.

This page intentionally left blank.

Acknowledgements

Reference style.

The copyright owners listed here have granted permission to reproduce material from their works.

The Source Notes lists other sources of quotations and photographs.

MITRE ATT&CK® Matrix, MITRE Corporation. https://attack.mitre.org/.

The Pyramid of Pain. SANS Institute, David Bianco. 2013. http://detect-respond.blogspot.com/2013/03/the-pyramid-of-pain.html.

This page intentionally left blank.

Introduction

Effective analytic support enables the movement, processing, and analysis of data to build understanding of the network, detect adversary activity, and inform the overall defense of the network.

This publication contains five chapters and three appendixes:

- **Chapter 1**—discusses the fundamentals of analytic support to defensive cyberspace operations. It begins with an introduction to analytic support data. The chapter goes on to introduce analytics and analytic support functions and discusses cyber analytics in the operations process and employment of analytics in defensive cyberspace operations. The chapter concludes with a discussion of industry taxonomies for cyberspace analytics.
- **Chapter 2**—discusses planning and assessment for analytic support, including how to collect, transport, process, and analyze data. It then discusses determination of technical information requirements and development of the analytic scheme of maneuver.
- **Chapter 3**—introduces data engineering. It explains the importance of data engineering to defensive cyberspace operations, and discusses data processing functions, including extract, transform, and load; data use; data management; data storage; and data retention requirements.
- **Chapter 4**—outlines and explains one methodology for analytic development to transform collected data into knowledge. The chapter discusses analytic solutions and procedures through which data passes during analysis.
- **Chapter 5**—addresses the execution of analytic support. It begins with a discussion of the execution of the analytic scheme of maneuver, including the analysis cycle, employment considerations, and support relationships. The chapter concludes with a discussion of enduring analytic support.
- **Appendix A**—discusses analytic measures of performance and measures of effectiveness.
- **Appendix B**—discusses data mapping techniques to assess, prioritize, and map data to set conditions for integrating analytic techniques into defensive cyberspace operations.
- **Appendix C**—discusses analytic techniques, tools, and methods.

This page intentionally left blank.

Chapter 1

Analytic Support Fundamentals

This chapter discusses the fundamentals of analytic support to defensive cyberspace operations. It begins with an introduction to analytic support data. The chapter goes on to introduce analytics and analytic support functions and discusses cyber analytics in the operations process and employment of analytics in defensive cyberspace operations. The chapter concludes with a discussion of industry taxonomies for cyberspace analytics.

INTRODUCTION TO ANALYTIC SUPPORT

1-1. Analytic support is the grouping of interrelated tasks that applies concepts from mathematics, computer science, and the cyberspace domain to derive insight from technical data. This insight comes from automating existing analyst tasks and providing analysts with new methods, processes, and capabilities for analysis. These insights support the two goals of data analytics—enhance local defense and enhance real-time situational awareness. Analytic support is rooted in the data and analytics leveraged, and it consists of three interrelated functions with continual planning and assessment cross-cutting all functions. Figure 1-1 illustrates the functions of analytic support.

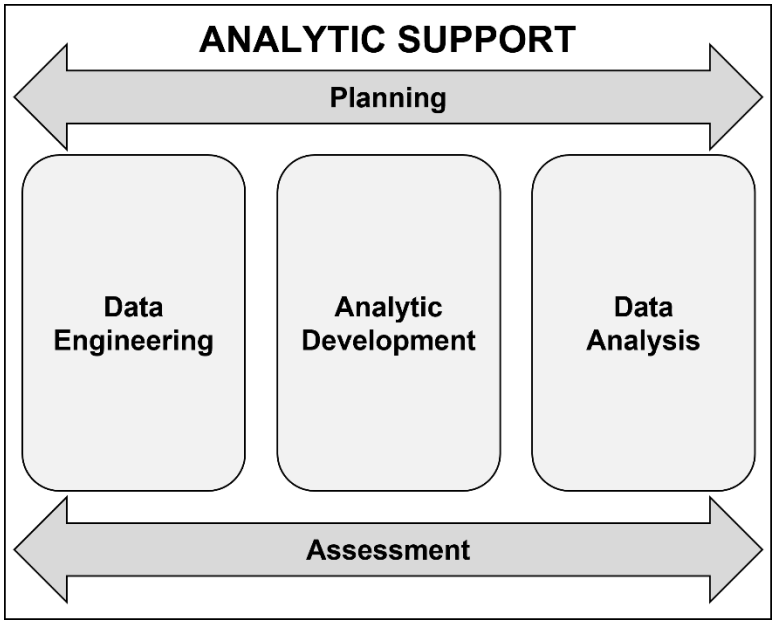

Figure 1-1. Analytic support functions

DATA

1-2. Analytics run on data, making data both a crucial precondition for analytic support and key determinant in its effectiveness. Through data collection and processing, data engineering ensures access to quality data. Since access to quality data is necessary to generate meaningful insights, data engineering is a key component of successful defensive cyberspace operations.

CATEGORIES OF CYBERSPACE DATA

1-3. While data requirements vary depending on the environment and mission, data used in cyberspace operations can generally be characterized as either host data or network data. Host data relates to the state or activity of a specific system—it includes application logs and running processes. Network data relates to the configuration and activity of the network—it includes firewall logs and network traffic captures. In many

instances, complete understanding requires both host and network data. This is particularly true for networking devices and network infrastructure.

DATA COLLECTION

1-4. Data collection ensures analysts have access to the data they need. Access to the right data directly correlates with analyst visibility and the likelihood of mission success. Analysts cannot answer analytic questions about data they cannot view or analyze. To have historic data for analysis, there must be persistent, broad-scale logging on the network. Data access during missions involves synthesizing the existing, but potentially disparate, data sources for retroactive analysis while tuning data collection toward prospective requirements. Often, data will be stored in a centralized data analysis platform, but this will vary based on mission considerations. For example, the volume of data may make transferring all data to a centralized platform impractical. In these cases, analysts must access the data on the devices—sometimes at the tactical edge where the data is being stored.

DATA PROCESSING

1-5. Data's utility is based on its quality and usability. There are many measures of quality, but quality data should be complete, accurate, relevant, consistent, and unique. For instance, a data source may be missing data fields or data from a time window, making it incomplete and reducing its quality. Further, different data sources may use different time zones, requiring modification or refinement to correlate with other data sets. Data processing can mitigate quality and usability issues by cleaning, deduplicating, normalizing, and using inference. Data can be enriched by adding additional information, such as geographic location for Internet Protocol addresses. Data processing can also reveal data gaps that will need to be addressed through analysis, especially if the gaps affect data sources that critical to the analytic scheme of maneuver. Data engineering and analysis can help mitigate these gaps.

ANALYTICS

1-6. *Analytics* are the systemic processing and manipulation of data to uncover patterns, relationships between data, historical trends, and attempts at predictions of future behaviors and events (NIST SP 1500-1). Analytics includes processing to turn data into information, and analysis to further turn that information into knowledge. Processing and analysis rely on software tools and automation. Data analytics is the use of the automation for analysis. It turns data into information. Because manual processing and analysis may require extensive time and expertise, data analytics significantly improve the efficiency and effectiveness of a single analyst. Figure 1-2 shows how processing turns data into information and subsequent analysis transforms information into knowledge. Commanders and staffs then apply judgment to transform knowledge into understanding.

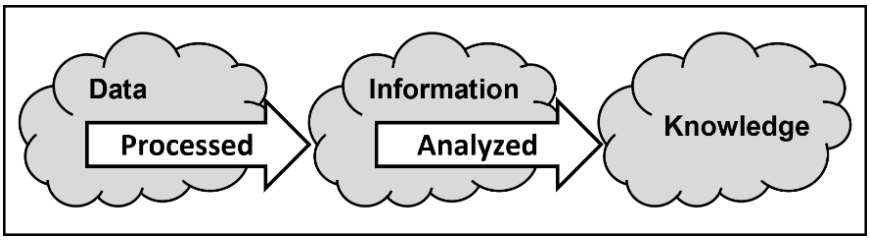

Figure 1-2. Data, information, and knowledge

1-7. There are four categories of analytics: descriptive, diagnostic, predictive and prescriptive. Descriptive analytics explain what happened. Diagnostic analytics explain why something happened. Predictive analytics forecast what might happen in the future. Lastly, prescriptive analytics recommend actions based on optimization and forecasting information. Prescriptive analytics are the most complex type of analytic solution to develop and use, but also generally provide the most value. Figure 1-3 on page 3 shows how the various types of analytic solution relate to one another in terms of complexity and value.

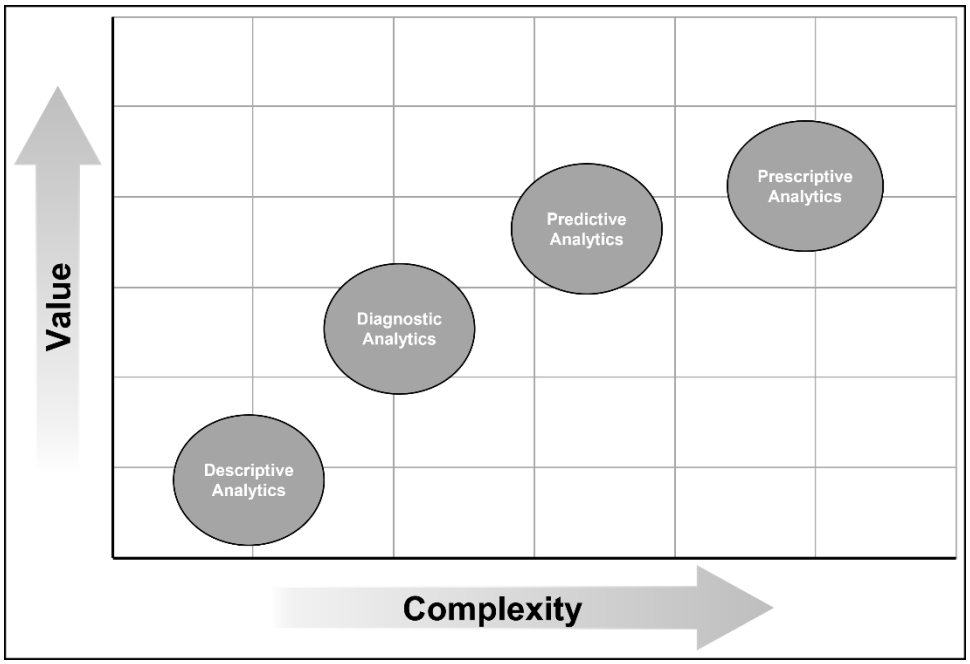

Figure 1-3. Analytic complexity by type

ANALYTIC SUPPORT FUNCTIONS

1-8. Analytic support consists of four primary functions that enable defensive cyberspace operations across echelons and mission. The analytic support functions are—

- Planning and assessment.
- Data engineering.
- Analytic development.
- Data analysis.

PLANNING AND ASSESSMENT

1-9. The process of analytic planning and assessment is continuous. The analytic support officer typically leads the process and integrates analytic planning and assessment into the operations process to identify how and where to apply data analytics. Planning and assessment both informs, and is informed by, intelligence preparation of the operational environment through refined understanding of the terrain, enemy, and friendly forces. Planning and assessment is also nested with the commander's critical information requirements. Although the initial analytic scheme of maneuver is generated in the planning phase, it continues to adapt throughout the operation through deliberate, iterative intelligence preparation of the operational environment and analytic scheme of maneuver generation. Assessments of analytic measures of performance and effectiveness inform this process and that of future missions.

DATA ENGINEERING

1-10. Data engineering sets the conditions for analytics to be applied against data, ensuring data access and quality. Data engineering includes transport; the extract, transform and load process; and data validation. The data engineer has primary responsibility for data engineering and works closely with the analytic support officer and the rest of the support element.

ANALYTIC DEVELOPMENT

1-11. Analytic development addresses gaps in data and analytics required to execute steady-state or demand-driven operations. Once an analytic solution is developed and tested, it is introduced into a singular analytic inventory for knowledge management and information sharing, providing a common understanding of analytic coverage.

DATA ANALYSIS

1-12. Data analysis is the process which generates findings and insights. Data analysis includes execution of the analytic scheme of maneuver and the analysis cycle. This function may include leveraging external enablers, such as analytic support cells, through support relationships. Figure 1-4 depicts the major interactions between the analytic support functions. Planning and assessment are informed by the other three functions and direct their execution.

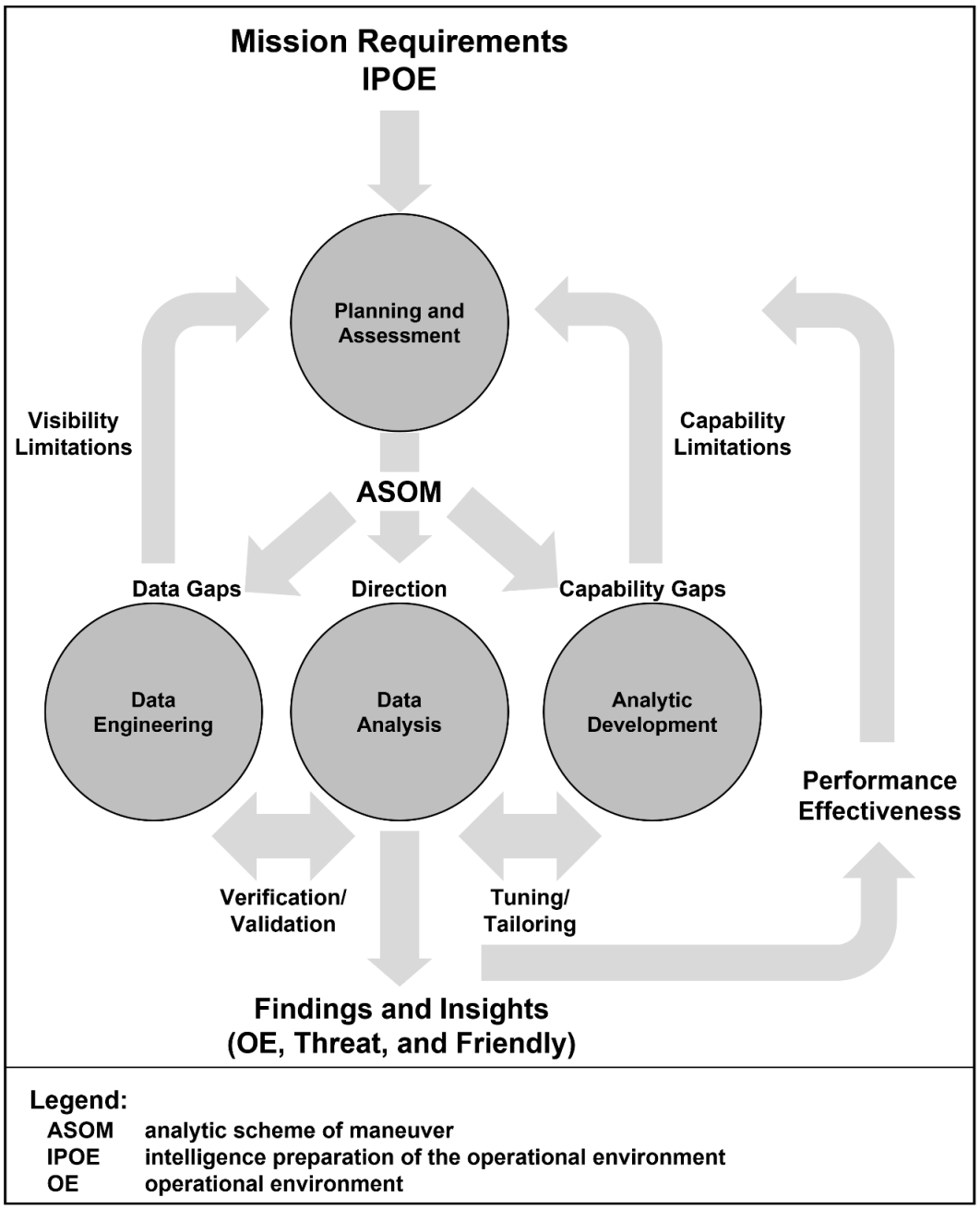

Figure 1-4. Interaction of analytic support functions

DATA ANALYTICS AND THE OPERATIONS PROCESS

1-13. Effective use of analytics requires their integration throughout the operations process. This applies in both hasty and deliberate operations. The unit leader is responsible for this integration, leveraging expertise from across the unit and staff. Analytic support officers advise and assist the commander in this process.

1-14. The use of analytics both informs and is informed by mission analysis, specifically intelligence preparation of the operational environment. For example, analyzing network data can define the software applications used in the operational environment, while understanding the network design may drive data collection strategies. Additionally, the commander's critical information requirements generated during mission analysis drive much of the analytic planning.

1-15. During the preparation phase, defensive cyber forces may use data analytics to develop situational understanding. They may also take steps to ensure they have the correct analytic capabilities and data. When data is not available, cyber elements employ their sensors to close data gaps.

1-16. As the mission progresses, analytics identify conditions for actions and actions set conditions for analytics. Units may change network configurations or initiate data collection to enable analytics. Friendly actions may, both incidentally and by design, create data that is observed and reported in analytics. Conversely, analytics may inform decisions to initiate planned or unplanned response actions. By integrating analytics into the courses of action during the military decision-making process, the commander and staff can exploit these opportunities while reducing potential friction.

1-17. Units can use analytics to assess measures of performance and measures of effectiveness and to help commanders decide on necessary follow-on operations. See appendix A for example measures of performance and measures of effectiveness for the integration of defensive cyberspace operations.

EMPLOYING ANALYTICS IN DEFENSIVE CYBERSPACE OPERATIONS

1-18. While the benefits of analytics and data science are frequently discussed, there is a less robust body of knowledge discussing how to employ them in tactical operations. At its core, employing analytics is about asking questions of data to produce understanding. Therefore, successful application depends on the questions asked, the data analyzed, and the analytic tools used for analysis. Units must select these carefully, based on their operational context.

ASKING THE RIGHT QUESTIONS

1-19. The operations process drives a unit's need to understand the operational environment, friendly forces, enemy forces, and their information requirements. Intelligence preparation of the operational environment provides the framework for this understanding and leads to information requirement development. The Army design methodology and military decision-making process, assist commanders and staffs in generating the commander's critical information requirements. Finally, additional requirements can arise during execution and assessment phases. From these information requirements, analysts identify specific questions to be answered using technical data.

IDENTIFYING THE RIGHT DATA

1-20. Mission planning techniques identify the data required to answer current technical questions and the data that may be required for future questions. Enemy, friendly, and terrain considerations shape what data is available and inform the decision regarding which data to use. Units select a combination of data sources that will provide the necessary information within mission constraints. Data processing and collection efforts enable the unit's access to the right data, while data verification and validation ensure the data is correct and complete.

IDENTIFYING THE RIGHT ANALYTICS

1-21. Analytics answer questions by transforming data into information and knowledge. A valid analytic may answer the question with data that can be made available. However, choosing the right analytic is mission-

specific and requires balancing a confluence of factors relating to the requirements and results of the analytic. These may include the—

- Mathematical assumptions and limitations of the analytic.
- Costs for development, training, tuning or deploying the analytic.
- Analyst burden.
- Robustness of the results.
- Confidence—including considerations of accuracy and recall—in the results.

1-22. Analysts synthesize situational understanding from many observations. This synthesis often requires analysts to use a collection or sequence of analytic tools.

HOW TO EMPLOY ANALYTICS AGAINST DATA

1-23. At a minimum, the data is made accessible to the analytic tool; the analytic software code runs to produce a result. However, except in trivial cases, routine applications, or pristine environments, analytic employment is neither linear nor a single, simple solution. Instead, it is an iterative process of development and analysis. Inputs must be validated, verified, and may require additional processing. The analytic support officer may need to adjust and refine the software coding and refinement, and the outputs of the analytic tool must be validated and contextualized. Additional analysis may be required to expound upon or corroborate findings.

AVAILABLE SUPPORT MECHANISMS

1-24. As defensive cyber forces keep pace with emerging threats, they need support mechanisms to enable analytic growth. Analytic support cells address the gap in analytic support at echelon. Analytic support includes planning and assessment, analytic development, data engineering, and data analysis.

INDUSTRY TAXONOMY

1-25. Adopting a common language and framework is essential to establish a knowledge base for cyber analytics. Common, widely recognized frameworks such as the MITRE adversarial tactics, techniques, and common knowledge (ATT&CK®) knowledge base bring different advantages aligned to mission requirements when hunting and tracking an adversary.

1-26. MITRE ATT&CK® is used for conveying individual adversary actions, relations to other actions, how these sequences of actions relate to tactical adversary objectives, and how the actions correlate with data sources, defenses, configurations, or other countermeasures. The MITRE ATT&CK® knowledge base is a robust, well-documented, and widely implemented framework that aids analytic support officers with an extensive library of detection mechanisms aligned to observed adversary tactics and techniques.

Chapter 2

Planning and Assessment

During analytic planning, the unit determines how to collect, transport, process, and analyze data. They must determine the technical information requirements, both from the commander's critical information requirements and broader requirements for situational understanding and decompose them into specific questions with associated data requirements and a method or tool for analysis. This process of decomposition and re-association is dictated by both general understanding of the cyberspace domain and the specific threat-informed situational understanding that results from intelligence preparation of the operational environment. The analytic scheme of maneuver is a result of analytic planning.

COMPONENTS OF THE ANALYTIC SCHEME OF MANEUVER

2-1. The analytic scheme of maneuver is the plan to collect and analyze technical data to meet specific information requirements. It identifies what data to analyze, how to analyze it, and why it is being analyzed. An analytic scheme of maneuver is like an information collection plan but also addresses the conduct of analysis. It helps drive analytic development, data collection and engineering, and analysis efforts. The time available, echelon, and the unit's experience influence the level of detail captured in the analytic scheme of maneuver. Less experienced units will require more detailed plans to be effective, while hasty planning or planning by higher echelons may preclude this level of detail. Refer to FM 3-55 for more information about information collection.

WHAT DATA TO ANALYZE

2-2. Generally, the analytic scheme of maneuver specifies the type of data and location. For log data, the type is identified by the source software and the log type. The location is typically based on the logical or physical layer of cyberspace but could be a dataset or data repository. The plan may use named areas of interest to concisely refer to data sources and locations.

HOW TO ANALYZE

2-3. The method of analysis may be an existing analytic capability, a standard technique, or brief description. It may also specify additional inputs such as thresholds or reference data.

WHY TO ANALYZE

2-4. To show the purpose of the analysis, the analytic scheme of maneuver must identify the nesting of indicators and evidence with specific information requirements. The specific information requirement is generally a commander's critical information requirement but may include more general requirements necessary to inform situational understanding.

ADDITIONAL INFORMATION

2-5. The analytic scheme of maneuver must include sufficient information to direct subordinate units, such as who performs the analysis, data transport and storage, the systems used for collection or analysis, and reporting or follow-on actions required. Units use tools such as data flow diagrams, collection overlays, and analysis matrixes to depict the plan. Figure 2-1 on page 8 depicts an example of a collection and analysis matrix as a portion of a notional analytic scheme of maneuver.

#	Indicator	#	Evidence	#	Data	NAI	Action
1.3	Has the adversary leveraged a WinRM vulnerability? (T1021.006)	1.3.1	Brute force attempts, excessive failed logins	1.3.1.1	Security EID4625 (failed logon)	1, 2, 3	• Excessive failed logons from a single source to a single destination host • Excessive failed logons from multiple hosts to a single destination host • Anything above paired with a successful login indicates likely compromise
				1.3.1.2	Security EID4648 (logon attempt per creds)		
				1.3.1.3	Security EID4740 (user locked out)		
				1.3.1.4	Security EID4782 (password hash accessed)		
		1.3.2	Deviation from normal pattern of execution	1.3.2.1	WinRM execution list	3	Get WinEvent-ListLog *winrm* | fl*, get last log time and put into inputs.conf
		1.3.3	Scripts initiated by WinRM	1.3.3.1	WinRM EID6 (remote handling activity started)	3	Investigate events outside of allow-listed scripts
				1.3.3.2	WinRM EID169 (remote handling activity authenticated)		
				1.3.3.3	WinRM EID142 (remote shell initiation attempt)		
				1.3.3.4	wmiprvse.exe child processes		
		1.3.4	Abnormal process creation	1.3.4.1	WinRM EID4688 (process created)	3	Establish baseline then differential
				1.3.4.2	Process list	3	Investigate non-allow-listed parent processes that spawn WinRM child processes

Legend:
EID event identification
NAI named area of interest
WinRM Windows Remote Management

Figure 2-1. Example collection and analysis matrix

DEVELOPING THE ANALYTIC SCHEME OF MANEUVER

2-6. The analytic scheme of maneuver is developed from commander's intent and desired end-state and starts with the commander's critical information requirements. Units develop commander's critical information requirements as part of mission planning. The analytic scheme of maneuver focuses on answering the commander's critical information requirements with technical data.

2-7. During analytic scheme of maneuver development, units refine the commander's critical information requirements and any other technical information requirements that aid in answering commander's critical information requirements or the broader mission. They decompose these information requirements into indicators, evidence, and actions. Figure 2-2 is an example of a priority intelligence requirement decomposing into evidence.

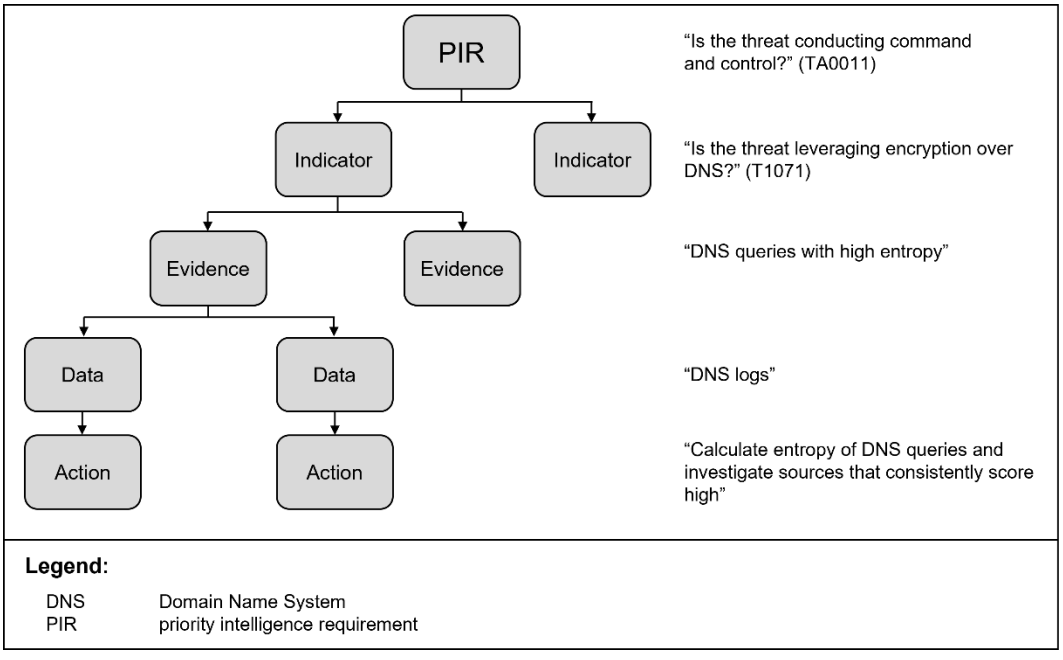

Figure 2-2. Analytic scheme of maneuver development

INDICATORS

2-8. In the context of analyzing cyberspace data, indicators are items of information that can support answering an information requirement. Indicators relating to an adversary's course of action can be broken down into multiple sources of information that contextualize attacker behavior. However, in the context of defensive cyberspace operations, indicators may also include information that informs understanding of friendly forces and the operational environment. In some contexts, it may be necessary to scope information requirements with intermediate requirements called essential elements of information before developing indicators. An example of a malicious indicator is a regular user conducting enumeration of a networking device by manually running the command line. In this example, the pieces of evidence are historical command line logging, account privilege schemes, network traffic evidence and an understanding of the user's normal responsibilities.

2-9. Statements of evidence lend strength to indicators. These are usually observations about the data that reveal patterns or actions on systems or networks. Evidence can range in complexity from the presence of simple signatures, such as Internet Protocol addresses or file hashes, through complex observations about patterns or actions.

THE PYRAMID OF PAIN

2-10. The Pyramid of Pain—developed by the SANS Institute—is useful when refining indicators of adversary activity. It depicts how certain evidence is more challenging to detect but may have larger impacts on the adversary. It additionally reveals that the evidence that is easier to detect is more fragile. That is, the evidence that is easier to collect is also easier for the adversary to change, breaking the detection. Adversary tactics, techniques, and procedures are preferred but the most challenging to detect. Figure 2-3 depicts the pyramid of pain.

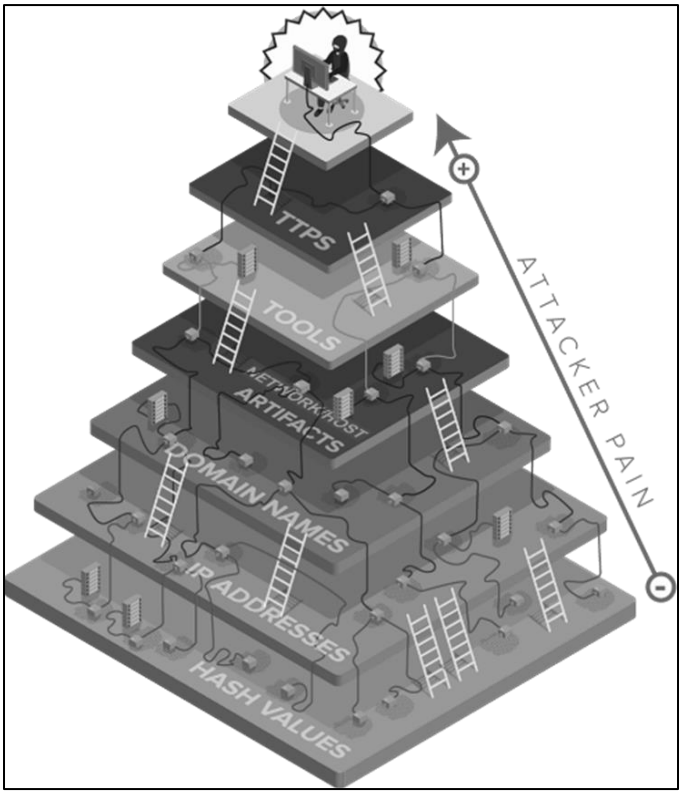

Figure 2-3. The pyramid of pain

REFINEMENT

2-11. By identifying indicators and evidence for information requirements, units refine the questions they are asking. The selected indicators and subsequent evidence dictate the specificity of the questions. Units must use sufficient specificity to guide both the preparation and execution. Achieving this specificity may require iterative refinement of the information requirements into indicators and evidence.

TRANSFORMING DATA INTO EVIDENCE

2-12. Input data, such as logs, configuration settings, and network traffic is processed and analyzed to produce results that may be evidence of an indicator. The processing and analysis of the data may include both manual and automated steps. In all cases, the data used constrains the applicable analysis techniques and the potential results. Selection of analysis methods and data sources is a matter of trade-offs, assessing the costs against the capabilities relative to efficiency and effectiveness.

ANALYTIC EMPLOYMENT CONSIDERATIONS

2-13. Deciding which analytic solutions to employ during defensive cyberspace operations is largely driven by intelligence preparation of the operational environment. Understanding the operational environment and enemy forces allows leaders to prioritize which, where and how analytic solutions are employed. Several significant factors for doing so are detailed in Table 2-1. By taking these into account, leaders can increase their unit's ability to answer commander's critical information requirements with increased confidence.

Table 2-1. Analytic employment considerations

Title	***Description***	***Example***
Mission Type	Tactical tasks have different objectives and analytic solutions must focus toward achieving these.	The goal of a hunt is to detect an adversary's presence and deny them their objectives. The goal of a mission thread defense is to ensure a network or system is capable of meeting commanders' requirements. The latter will focus on defending critical terrain and will employ analytics accordingly.
Preexisting Capabilities	Every network has existing defensive cyberspace operations capabilities which must be either de-conflicted or incorporated.	Network A has an endpoint security solution that detects Excel macro abuse. If these are determined to be adequate, defensive cyber forces should not prioritize employment of additional analytic solutions to detect Excel macro abuse.
Data Availability	If an analytic solution requires a level of logging that doesn't exist, don't employ it until that logging is enabled.	An analytic solution requires Domain Name System queries and responses to be logged. In Network A, only Domain Name System queries are logged. Therefore, additional logging must be enabled before this solution is employed.
Time to Tune	Analytic solutions that require minimal tuning should generally be employed first.	An analytic solution detects abnormal sever behavior but requires manual classification of all network objects. Although potentially effective, the tuning of this is time consuming and will require knowledge of the operational environment that usually comes later in an operation.
Probability of Enemy Detection	Select analytic solutions that have the highest likelihood of detecting enemy activity.	Intelligence indicates an adversary is utilizing domain name system tunneling for exfiltration. The analytic support officer should choose to prioritize employment of several domain name system tunneling analytic solutions before employing a solution for outbound secure shell protocol exfiltration.

Chapter 3

Data Engineering

This chapter introduces data engineering. It explains the importance of data engineering to defensive cyberspace operations, and discusses data processing functions, including extract, transform, and load; data use; data management; data storage; and data retention requirements.

DATA CONSIDERATIONS

3-1. Data engineering is a critical component of cyberspace operations. A cyber analyst's greatest chance of identifying malicious cyber activity lies within the data artifacts that are being collected and analyzed in the environment. Data engineers create and maintain data pipelines, develop scripts for data analysis and correlation, and assist with the development of the analytic scheme of maneuver. Data engineers work closely with analytic support officers and cells during operational and tactical planning to prioritize data sources and collection efforts. Data engineers ensure that data produced at lower echelons is extracted, transformed, and loaded to match directed data schemas.

3-2. Ideally, all data that is centrally aggregated can be queried using a common information model. Each data source structures data in different ways, but often shares some key information (for example, Internet Protocol address, or host name). Because the most effective analytic solutions use multiple data sources, key information needs to be easily queried. When data is not formatted in a common information model, computational and analyst time is spent conducting conditional checks on various permutations of a field name.

3-3. There are two ways data engineers traditionally develop a common information model to enable data queries. One way is to manipulate all the data as it enters the centralized location—typically a security incident and event management database—by modifying field names to a uniform structure. For example, all source Internet Protocol field names would be changed to src.ip. The other way data engineers might create a common information model is with aliases that relate similar field names. That way if an analyst queries src.ip, src_ip will also be queried. In addition to standardizing field names, data engineers may need to normalize the data into a common format. For example, some data sources will represent hardware addresses as AA:BB:CC:DD:EE:FF, while others will represent them as AA.BB.CC.DD.EE.FF. Standardizing the data format enables the use of automated analytic tools.

***Note*.** Data normalization is a process that facilitates a more cohesive form of data entry, essentially 'cleaning' the data. When an analyst normalizes a data set, they reorganize it to remove any unstructured or redundant data and enable a more logical means of storing that data.

3-4. At the tactical edge, data engineers work with analytic support officers, analytic support cells, and mission partners to collect and deliver identified data in the proper format to the applicable data analysis systems at the appropriate level. For example, data engineers will manage operational data analysis systems (such as security incident and event management, Splunk, or Elastic) and are responsible for ensuring operational data is available for analysis in the big data platform.

3-5. Data requirements are driven by the analytic scheme of maneuver. To leverage data for analysis, the data engineer is required to identify potential sources of data. Potential sources can include, but are not limited to, open-source intelligence data feeds, threat indicators of compromise feeds, network traffic, host traffic, and change logs. Data engineers, analytic support officers, and leadership work with mission owners to create a data collection plan based on the data required versus data available.

DATA PLATFORMS AND DATA TRANSPORT

3-6. Data engineering begins with identifying where the analysis will be applied to the data. Either the data is brought to the analytics, or the analytics are brought to the data. The simplest approach is to have all data collected in a central location, such as a database or security incident and event monitoring system.

3-7. Network latency, volume and security classification of data, distribution of resources, and other considerations often require a hybrid approach. Data may be pre-processed and even analyzed at or near the point of collection, with certain data or results being centralized. In conjunction with this determination, units engineer their data transport solution. The solution may involve network changes and must address what data is sent, how, and when, while preserving the integrity of data and consistency of the databases.

DATA PROCESSING

3-8. Once the correct data is collected and aggregated to a central location, data engineers are charged with making the data useful. Quality data should be complete, accurate, relevant, consistent, and unique. The extract, transform, load pipeline is how data engineers turn poor quality data into high quality data. While the extract, transform, load pipeline is explained as a linear process, situations often require additional extractions, transforms, or loads, depending on the data pipeline.

EXTRACT, TRANSFORM, LOAD PIPELINE

3-9. The extract, transform, load pipeline is a process where data is extracted from its original source, transformed (cleaned, sanitized, scrubbed), and loaded in a more useful format into an output data container. The data may be collated from multiple disparate sources and may be output (loaded) to one or more destinations. Extract, transform, load processing is typically automated using a software application.

Extract

3-10. Extract is the first step in the extract, transform, load pipeline. Data is gathered through numerous collection methods that consist of application programming interface calls, sensor collection, and intelligence feeds, all of which are formatted in their own respective schema.

Transform

3-11. Transform is the second step in the extract, transform, load pipeline. The data engineer normalizes raw data using various techniques to standardize, deduplicate, and validate the data. Columnizing data is the process of flattening hierarchical data into a simple tabular format that an analyst can test against. During the transformation pipeline the data is stripped of any columns which are deemed uninteresting or irrelevant by the mission owner. During the extract, transform, load pipeline build phase, the data engineer will work with the network engineers or technicians to identify the most efficient transport of data to avoid bottlenecks and potential sources of congestion. The data engineer also coordinates with the team to identify and emplace security controls designed to protect data integrity within the pipeline.

Load

3-12. Load is the last phase in the extract, transform, load pipeline. The data engineer ensures data is optimally loaded after extraction and transformation. There are several considerations based on the type of loading process and whether the data will terminate at a data warehouse or security incident and event management platform. Storage and format of the data is determined based on data access requirements.

DATA VERIFICATION AND VALIDATION

3-13. Through data processing, data engineers verify and validate the data. They must ensure collection is complete in terms of coverage, data sources, and content. They must also ensure the data is correct. Data transformations must be applied properly. Nonsensical data, such as dates in the future, indicate potential problems with collection, processing, and understanding. Data must also be consistent—reuse of Internet Protocol addresses across networks and lack of time synchronization are common consistency challenges.

DATA MANAGEMENT

3-14. Managing the storage of data is a key responsibility of a data engineer. They must design a system that balances the availability of data with the inherent cost constraint that comes with storing large amounts of data. Mission requirements may further dictate other characteristics of the data storage solution design. Data management includes managing the use, storage, and retention of data.

DATA USE

3-15. The data engineer evaluates the use of data provided to inform improvements through all stages of the data pipeline. This evaluation can be performed using various methods to include integrating measures of performance and measures of effectiveness or by MITRE ATT&CK® tactics, techniques, and procedures coverage percentage.

DATA STORAGE

3-16. Data storage uses a system of hot, warm, and cold storage to manage data in a cost efficient but available manner. The requirement to preserve data in its original state must be balanced with storage requirements for transformed data that is useful to analysts. Additionally, units must consider the computational cost and time required to query large amounts of data and select storage solutions suited to their anticipated use cases.

DATA RETENTION

3-17. Retention policies are important to ensure data is compliant with applicable laws, regulations, and standards. Data should be stored for as long as it is required, and eventually destroyed when it is not. Redundant array of independent disks configurations are good backup solutions to ensure the availability of the data when needed.

This page intentionally left blank.

Chapter 4

Analytic Development

Analytic solutions are developed and employed to transform collected data into knowledge. For this discussion, an analytic solution refers to a tailored query, piece of code, or a set of procedures through which data passes during analysis. Using a standardized procedure to plan, develop and employ analytics enables consistent quality. This chapter outlines and explains one methodology for analytic development.

GAP ANALYSIS AND REQUIREMENT DEVELOPMENT

4-1. Through planning and assessment, the team may identify gaps in data or in analytic capabilities. In some instances, these gaps can be resolved with other data or capabilities. Data engineers become responsible for the data gaps and analytic support officers are responsible for analytic (capability) gaps that are identified in the planning process. Capability gaps for crucial indicators or evidence often suggest a commander's critical information requirement cannot be answered, and therefore the unit cannot effectively complete the mission.

4-2. Resolving analytic capability gaps is a core function of an analytic support officer. This function occurs during all phases of an operation, including recovery and training phases. The analytic support officers work with other cyber professionals to develop analytical methods and analytic solutions to fill the existing capability gaps. While network and host analysts are expected to develop queries using simple methodologies, analytic support officers are expected to apply advanced statistical analytic methods to one or multiple datasets. These advanced analytic solutions are geared at identifying anomalous and malicious cyber activity or understanding the operational environment.

4-3. When the unit assigned to the mission is unable to resolve a capability gap, the analytic support officer is responsible for providing the technical information necessary for a higher echelon or external organization to work on resolving the gap. While an analytic support officer can serve as the point of contact for this coordination, it will generally serve the mission element best if the support element lead manages it and works with the analytic support officer when necessary.

COLLECTION AND ANALYTIC MANAGEMENT

4-4. Collection and capability management are the processes of ensuring that analysts possess the appropriate data, tools, and platforms to effectively execute defensive cyberspace operations and can answer commander's critical information requirements.

4-5. Identification of gaps in data coverage during mission planning processes will drive collection efforts required to apply data analytics. Mapping data sources to industry-aligned tactics and techniques using MITRE's ATT&CK® knowledge base is one method to understand data completeness where weight criteria can be applied to data sets collected. An effective visual representation of data completeness is overlaying a heat map on the MITRE ATT&CK matrix to show which technique have greater or lesser visibility. This data mapping technique allows commanders to see what is being logged within the terrain. See appendix B for a discussion of data mapping techniques.

4-6. Once gaps in data coverage are identified, collection plans focus on efforts to gather necessary data sets to answer commander's critical information requirements. There are various methodologies that can be applied when addressing data coverage gaps such as integrating concepts, such as high-payoff data sources that give analysts larger adversary technique coverage with just a few data sources. Additionally, data collection efforts may be prioritized based on prioritization of commander's critical information requirements. Other collection considerations include ease of collection, resource requirements, network impact, and data overlap. Figure 4-1 on page 16 shows an example of using the MITRE enterprise ATT&CK® knowledge base to identify data sources that provide the greatest coverage of adversary tactics, techniques, and procedures. This methodology is tailorable to the specific tactics, techniques, and procedures of a particular threat actor to identify which data sets analysts need for a given mission.

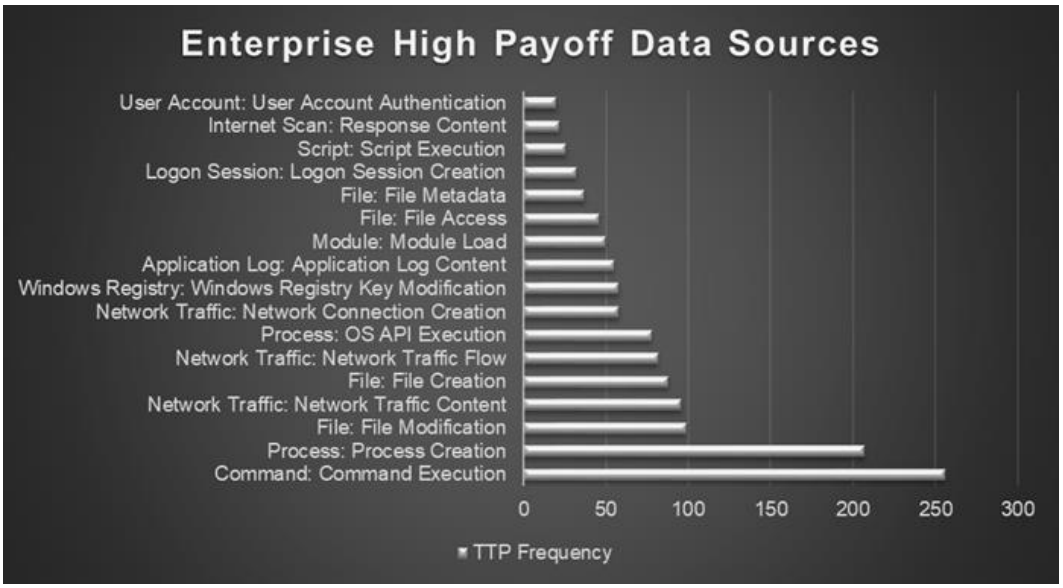

Figure 4-1. Example high-payoff data sources

4-7. Analytic management plays a key role in analytic support planning and preparation for execution of data analytics in defensive cyberspace operations. Management of analytics includes identification of readily available analytic solutions and prioritization of analytic gaps to be addressed during analytic development. Mapping methodologies can be similarly applied to understand analytical capabilities aligned to a common industry standard. One analytic prioritization methodology is to focus on the most common adversary tactics, techniques, and procedures across multiple threat actors. Figure 4-2 shows an example of prioritization based on the most common adversary tactics, techniques, and procedures.

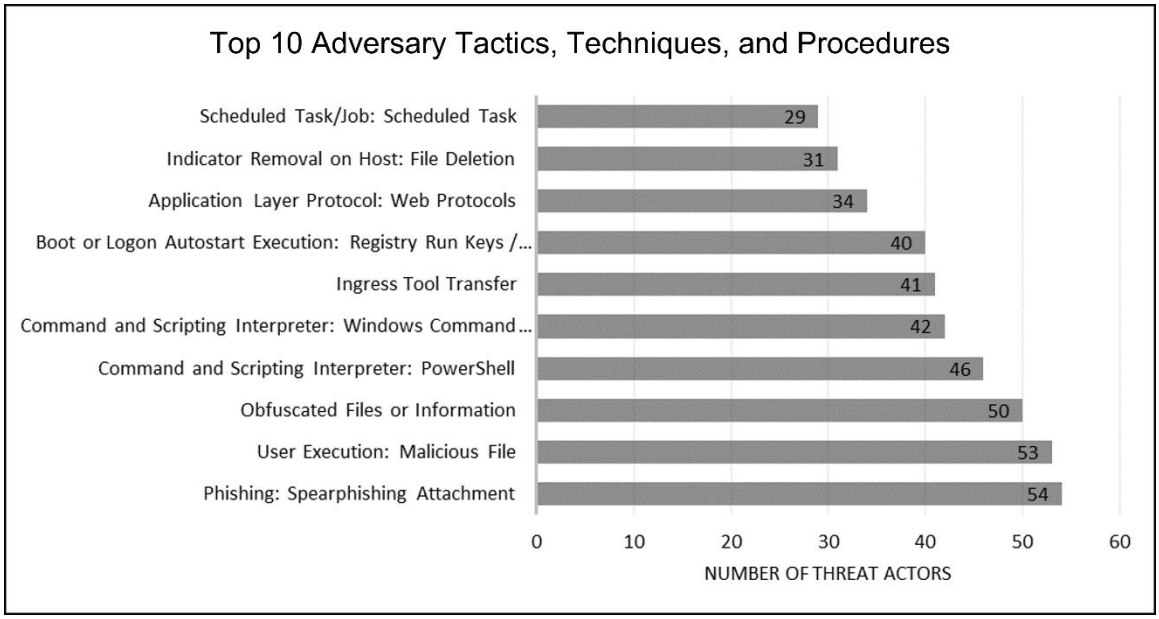

Figure 4-2. Most common adversary tactics, techniques, and procedures

ANALYTIC DEVELOPMENT

4-8. The integrated data analytic approach (see figure 4-3 on page 17) represents a data science methodology to guide development of analytic solutions that answer information requirements. The process begins during intelligence preparation of the operational environment where information and analytic gaps are determined.

From these, commanders prioritize analytic development and define which operational requirements to action. Once defined, an analytic support officer begins discovery by conducting data exploration to validate the data's integrity, determine relevant variables, and eliminate extraneous information. The analytic support officer develops analytic solutions using simulated data to prototype solutions for further refinement. Through the final iterative development operations process of testing and refinement, the analytic support officer produces a viable analytic solution that can be deployed into an environment.

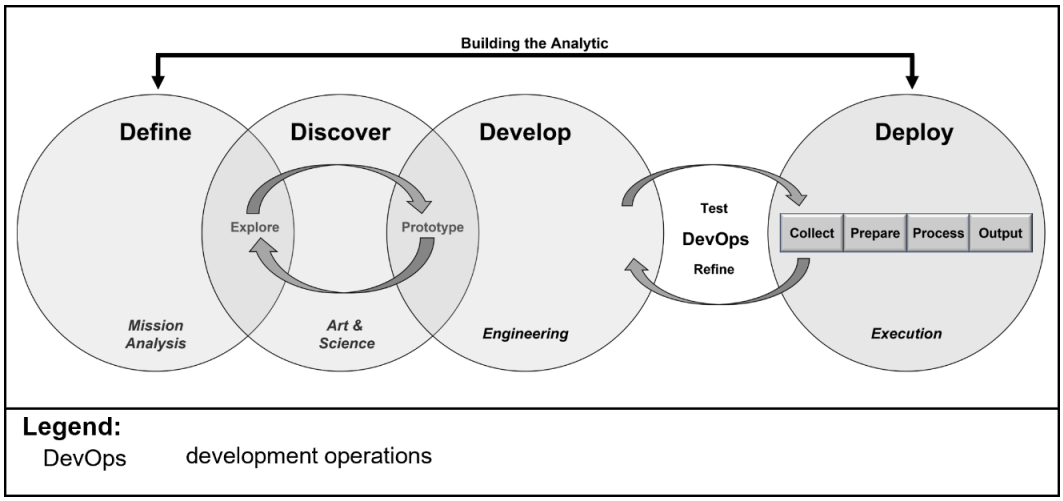

Figure 4-3. Integrated data analytic approach

4-9. Incorporating cyber threat emulation elements into testing and refining allows for a safe way to validate analytic accuracy. Before doing so, a testing environment that matches the target environment's structure must be developed. Although outside the scope of this document, an ideal testing environment accounts for several factors from the target environments to include user actions, admin actions, network traffic, abnormal-but-benign behavior, and realistic sensor placement. Once built, the analytic solution is deployed, and cyber threat emulation can emulate techniques which should trigger the developed solution. Cyber threat emulation can include full campaigns that emulate specific adversaries or atomic tests which only simulate a singular technique. By doing this, analytic support officers can increase the number of malicious events that are properly identified and optimize analyst efforts.

4-10. In defensive cyberspace operations, the testing process is of particular importance due to the high degree of accuracy required. The preponderance of analytic solutions in defensive cyberspace operations focus on detecting outlier events and classifying them as either benign or malicious. Due to the magnitude of data generated from most networks, a solution with a fractional error rate can result in potentially thousands of errors daily. Table 4-1 outlines the different types of these errors. Benign classification is preferred over malicious classification to prevent excessive false positives, which may overload analysts' cognitive capacity. However, this may result in a higher number of false negative classifications. The key is to be aware of the risks associated with a given scenario, implement mitigation factors that reduce risk to an acceptable level, and continue the mission. The testing process, and the time spent testing, allow an analytic support officer to tailor the analytic solution to the target environment and reduce classification errors.

Table 4-1. Binary classification outcomes

Type	***Description***
True Positive	When a malicious event is classified as malicious.
True Negative	When a benign event is classified as benign.
False Positive	When a benign event is classified as malicious.
False Negative	When a malicious event is classified as benign.

4-11. Although every environment is different, many analytic solutions can be modified and reused across different mission sets. Although short-term effectiveness is gained in rapid development, deliberately designing for reuse, and sharing results in long-term yield. Therefore, it is important to allocate time during

analytic development to do so. This means creating analytic solutions that are easily migrated to various analysis platforms and account for datasets with nonuniform field names. Examples of this include digesting threat intelligence in structured threat information expression, defining analytics in open signature format Sigma, and handling data in the common information model. See appendix C for detailed information on analytic techniques.

Chapter 5

Analytic Support Execution

This chapter addresses the execution of analytic support. It begins with a discussion of the execution of the analytic scheme of maneuver, including the analysis cycle, employment considerations, and support relationships. The chapter concludes with a discussion of enduring analytic support.

EXECUTING THE ANALYTIC SCHEME OF MANEUVER

5-1. The analytic scheme of maneuver is a living document that should be built iteratively and constantly refined and updated. During the initial planning and development stages, the analytic scheme of maneuver should reflect a threat template or doctrinal template—a combination of known or anticipated threat tactics, techniques, and procedures absent mission or terrain specifications. As planning elements conduct mission and terrain analysis, the analytic scheme of maneuver grows into a mission-specific document that reflects a situation template—hand-selected adversary tactics, techniques, and procedures based on the mission, terrain, and relevant intelligence. The analytic scheme of maneuver undergoes continual refinement as continued planning and initial survey occur to reflect further situational understanding of the network and the threat.

5-2. Through the analytic scheme of maneuver, the analytic support officer supports the mission lead as they determine the timing and tempo of the mission. In nearly all instances, there are more tactics, techniques, and procedures relevant to the mission than there are resources (time and analysts). Thus, the analytic approach to a mission must be deliberate and prioritized. The analytic support officer communicates, in aggregate, the information pertinent to the mission lead's ability to make decisions regarding the timing, tempo, and disposition of their forces. Considerations for prioritizing the analytic scheme of maneuver include relevant intelligence, key terrain, mission-relevant terrain, and the commander's desired end state.

DATA EXPLORATION

5-3. Units explore their data to gain general familiarity with underlying datasets, uncover preliminary patterns and identify points of interest. They validate that the expected data is collected and analyze any unexpected data. They verify that data processing is occurring properly. Data exploration occurs at the beginning of the operation and recurringly throughout the operation. in order to gain greater understanding of the dataset. While conducting data exploration, analysts focus on understanding four major factors shown in table 5-1.

Table 5-1. Defensive cyberspace operations data exploration factors

Factor	***Description***
Producer	Understanding what generated the event described by the data. This considers the underlying technologies and the actions being described by the data.
Collector	Understanding what collected the data and how it was collected. This considers the limitation of the collector or sensor and what data may not exist because of it
Structure	Understanding the structure of the data collected. This considers the layout of important fields, naming conventions and unique nomenclature
Completeness	Understanding how complete the data in relation to the area of operations and actions of interest. This considers if all transaction information was captured, the timeframe the data covers and relevant gaps over said timeframe

ANALYSIS PROCESS

5-4. *Analysis* is the compilation, filtering, and detailed evaluation of information to develop knowledge or conclusions (ATP 2-33.4). Analysis assists in understanding the operational environment, identifying adversary activity, and inform decision making. The analytic scheme of maneuver provides overarching guidance on where, how, and why to conduct analysis. Units also conduct less structured analysis during data exploration. In both cases, analysis relies on both critical and creative thinking. Analysis often requires intuition as analysts seek meaning in the data and individuals develop their own analysis processes based on common components.

5-5. Analysis is always conducted in support of a higher objective. However, existing data and tools cannot answer all questions. In those cases, units develop new capabilities, collect additional data, make other modifications to minimize the impact of the gap, or describe and bound the information gap so that leaders can make prudent risk decisions.

Querying and Manipulating Data

5-6. Analysts and an analytic code apply a variety of techniques, to include the structured techniques discussed in appendix A. The techniques may be narrowly focused to answer specific questions or more broadly focused to identify trends, patterns, anomalies, or other items of interest.

Synthesizing and Contextualizing Results

5-7. As analysis yields results, those result are put into context to include other finding, the network context, the broader operational context. Contextual information might include software policies, workflows, and contemporaneous events of the organization that uses the network.

Evaluating and Reporting Results

5-8. As analysts generate and contextualize results, they may answer or support information requirements as evidence or indicators. In other cases, results may be inconclusive or lacking critical details. They may suggest hypotheses that must be confirmed. Results that meet the reporting thresholds are communicated for broader use. This communication includes intelligence reports that are used to disseminate threat signatures and characteristic.

Refining Questions and Methods

5-9. Some results are not initially useful, instead analysts should ask more questions. Analysts may develop hypotheses to explain results and then must test these hypotheses by asking additional questions and gathering evidence. In other cases, the results might be too broad, imprecise, or inaccurate and the analyst will ask additional questions or refine their methodologies and tools. For instance, a threshold value for an analytic may yield too many false positives to process and a threshold value or underlying model might need to be adjusted.

Visualization

5-10. Effective visualizations assist in generating and communicating insights. Visualizations include pie-charts, timelines, graphs, histograms, and other graphics that depict information visually. Figure 5-1 on page 21 shows the components of the analysis process.

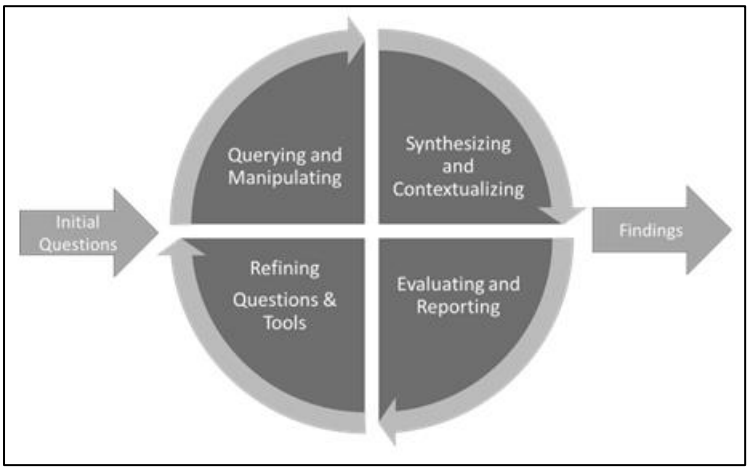

Figure 5-1. Components of the analysis process

SUPPORT RELATIONSHIPS

5-11. Commanders integrate analytic support into operations to concentrate efforts at the most advantageous place and time. When planning for integration, Commanders and staffs consider the fire support memory aid AWIFM-N (see table 5-2). Commanders then direct their analytic support cell to provide prioritized support.

Table 5-2. Fire support execution principles

Principle	*Description*
A	Adequate support for committed units.
W	Weight to the main effort or decisive operation.
I	Immediately available support for the commander.
F	Facilitate future operations.
M	Maximum feasible centralized control.
N	Never place analytic support in reserve

ENDURING SUPPORT

Analytic support tasks continue before and after missions. These tasks primarily focus on readiness and shaping conditions for future operations. The major analytic support tasks are—

- Training.
- Capability posturing.
- Knowledge management.
- Enduring priorities.

TRAINING

5-12. Analytic support and data analytics should be incorporated into both collective and individual training programs and include both analytic support personnel (analytic support officers and data engineers) and other members of the unit. Training should address the resources available, data, intelligence, software, concepts and methodologies, and the procedures to leverage them in support of operations.

CAPABILITY POSTURING

5-13. Units prepare for future operations by posturing their capabilities. A unit's analytic capabilities include the data, platforms, tools, and techniques available to the analysts. As campaigns progress, the threat evolves, and technologies change, new capabilities are required, and existing capabilities must be maintained. Posturing requires coordination with local network personnel, intelligence sections, and software developers and includes—

- Maintaining existing analysis tools, data collection, and analytic platforms.

- Refining employment, functioning, and documentation of existing capabilities.
- Developing new analytic methods and tools.
- Shaping collection in possible areas of operations.
- Integrating into network and intelligence reporting chains for tipping and cueing.
- Understanding the format and information provided by likely data sources.
- Pre-coordinating for physical and logical access.
- Setting conditions to scale data collection, transport, storage, and processing (compute).

KNOWLEDGE MANAGEMENT

5-14. For analytic support, knowledge management focuses on the tools and data available; the tactics, techniques, and procedures and standard operating procedures for using them; and the findings that result. It includes ensuring findings can be easily updated based on lessons learned from inside and outside of the organization. Intelligence personnel must be involved in managing findings, reporting, and understanding of the threat.

ENDURING PRIORITIES

5-15. Units, especially those with dedicated analytic support cells, may establish enduring priorities for data analytics. These are often related to situational understanding and tied to enduring missions or based on contingency planning. Personnel might maintain understanding of a network or operational environment, using analytics to understand the network composition and functioning. Alternatively, units may wish to conduct data analytics beyond the scope of the original mission, expanding the depth, breadth, or time of analysis based on the possibility that continued analysis will yield additional insights.

Appendix A

Analytic Measures of Performance and Measures of Effectiveness

Measures of performance and measures of effectiveness applied to analytics is an operational methodology to impose accountability metrics on data, analytics, and analysis processes. These accountability metrics shine a light on inefficiencies and enable defensive cyber forces to continue to grow against ever-evolving threats and technology. Typically, measures of performance and measures of effectiveness vary, depending on the analytic or intended mission end state. While current methods for employing measures of performance and measures of effectiveness are useful, automation of these functions in future analysis platforms will increase efficiency.

OVERVIEW

A-1. A *measure of performance* is an indicator used to measure a friendly action that is tied to measuring task accomplishment (JP 5-0) Applying measures of performance throughout defensive cyberspace operations enables leaders to evaluate analytic employment to answer the commander's critical information requirements. A *measure of effectiveness* is an indicator used to measure a current system state, with change indicated by comparing multiple observations over time (JP 5-0). Defensive cyberspace operations measures of effectiveness provide leaders an understanding of improvements or degradation of capabilities, data visibility, or efficiencies.

ANALYTIC MEASURES OF PERFORMANCE

A-2. Analytic measures of performance serve as a point in time assessment of analytic actions applied to answer commander’s critical information requirements. These measures enable leaders to assess if they have the right data and the right analytics to detect priority intelligence requirement-driven techniques. The following are example measures of performance applied when assessing defensive cyberspace operations analytic applications:

- Measure of performance 1: Analyzed data sources.
- Measure of performance 2: Analyzed techniques.
- Measure of performance 3: Analytics employed required to answer priority intelligence requirements.

ANALYTIC MEASURES OF EFFECTIVENESS

A-3. Analytic measures of effectiveness aid leaders in identifying efficiency statuses and capability gaps throughout defensive cyberspace operations. The following are example measures of effectiveness applied when assessing defensive cyberspace operations analytic applications.

- Measure of effectiveness 1: Time to identify, acquire, and refine data.
- Measure of effectiveness 2: Scope of data visibility.
- Measure of effectiveness 3: Time to detect techniques.

This page intentionally left blank.

Appendix B

Data Mapping Techniques

Understanding how data can be assessed, prioritized, and mapped is essential to set conditions for analytic techniques integration into defensive cyberspace operations. These techniques provide analytic support officers, data engineers, and leaders the ability to visually see and measure data collection efforts mapped to adversary techniques. Further development of data mapping platform integration is necessary to create a shared implementation across the defensive cyber force.

OVERVIEW

B-1. The MITRE ATT&CK® knowledge base is an industry standard for analytics and hunt operations and is a valuable resource for understanding threat-focused attack surfaces. One use of the MITRE ATT&CK® knowledge base is to associate tactics, techniques, and data sources. Tactics represent the why behind the employment of certain techniques and how the adversary achieves a tactical goal by performing an action. One essential component of detecting techniques is the required data sources including data components. Understanding how the framework is designed, what data sources are required to identify adversary techniques, and how that knowledge can be employed allows analytic support officers to make informed decisions during mission planning. Approaching the framework from a programmatic perspective allows analytic support officers and data engineers to understand what data is required to answer priority intelligence requirements.

MITRE ATT&CK® NAVIGATOR

B-2. The MITRE ATT&CK® Navigator is a web-based tool used for annotating and exploring ATT&CK® matrixes. It can visualize defense coverage, red and blue team planning, the frequency of detected techniques, and more. One of the inherent abilities of the tool is to visualize coverage and map frequency of techniques or data sources.

***Note*.** ATT&CK® Navigator is hosted on a public-facing website. There are no assurances of the confidentiality of any content uploaded. Units must consider operations security before uploading any content. Laws, regulations, and policy may impose additional restrictions on data uploaded. Organizations can host their own instances of ATT&CK® Navigator in compliance with the appropriate policies and regulations.

DEFENSIVE CYBERSPACE OPERATIONS INTEGRATION

B-3. This capability integrates into defensive cyberspace operations through—

- Web user interface.
- JavaScript object notation upload.
- Stand-alone Excel integration.

B-4. Through the web user interface, analysts can manually create layers tailored towards their specific operation or threat actor. Analysts can also generate JavaScript object notation files to load into the ATT&CK® Navigator and cross-reference collection plans against ATT&CK® data sources aligned to techniques. Finally, many of the ATT&CK® heat mapping capabilities are translatable into standalone Excel formulas mapped to an integrated visualization. These are a few ways of integrating MITRE's ATT&CK® mapping capabilities, but further integration into big data platforms and defensive cyber force equipment sets still needs to be explored.

ASSESSING COLLECTION EFFORTS

B-5. An advantage of using the MITRE ATT&CK® Navigator is the ability to score each technique based on the frequency of the data source and defensive cyber forces' ability to collect or analyze that technique. Applying this methodology during mission planning can be as simple as applying true or false statements aligned to collected data types as seen in figure B-1.

Network Traffic: Network Traffic Flow	TRUE
File: File Creation	FALSE
Network Traffic: Network Traffic Content	FALSE
File: File Modification	FALSE
Process: Process Creation	FALSE
Command: Command Execution	FALSE

Figure B-1. Example data source collection assessment

B-6. Once analysts define what data is being collected, generation of a heat map following the example in figure B-2 enables leaders to visualize technique coverage or gaps required to answer priority intelligence requirements.

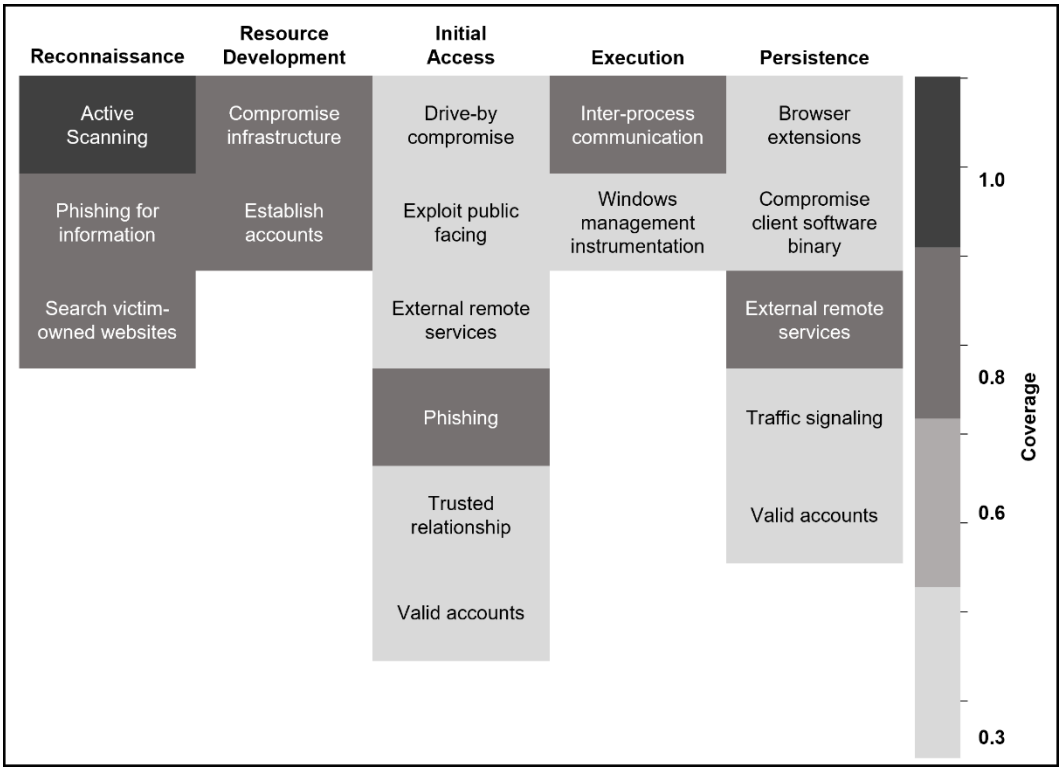

Figure B-2. Example heat map

IDENTIFYING HIGH-PAYOFF DATA SOURCES

B-7. Using the MITRE ATT&CK® knowledge base data, analytic support officers and data engineers can assess what data provides the most technique coverage. Additionally, this analysis informs data strategies when constrained by factors such as cost, capability, or storage limitations. To conduct this analysis, analytic support officers and data engineers apply a simple count by data source mapped to technique. This methodology is typically employed when defensive cyber forces are uncertain of the composition and disposition of enemy forces in their area of operations, and information concerning terrain is vague.

MAPPING THREAT ACTOR TECHNIQUES

B-8. Using the MITRE ATT&CK® Navigator web user interface, analysts can easily load and overlay known threat actors' techniques to visualize most likely or most dangerous scenarios. Additionally, during operations where there is a relatively thorough understanding of the operational environment to include detected threat techniques. The navigator could enable analysts to battle track adversary techniques. Further integration and testing in defensive cyber forces big data platform and equipment sets is needed to enable future automation capabilities based on alert or event detections.

This page intentionally left blank.

Appendix C

Analytic Techniques

Analysts, whether through manual analysis or by deploying automated solutions, use cognitive processes and analytic techniques and tools to generate knowledge and limit analytical errors. This appendix discusses two types of analytic techniques—cognitive and mathematical.

TECHNIQUES, TOOLS, AND METHODS

C-1. Analysts use cognitive processes and analytic techniques and tools to solve problems and limit analytical errors. They complete these processes using a variety of analytic techniques, tools, and methods.

- **Technique**—a way of doing something by using a special knowledge or skill. An analytic technique is a way of looking at a problem, which results in a conclusion, an assessment, or both. An analytic technique usually guides analysts in thinking about a problem instead of providing a definitive answer as typically expected from a method.
- **Tool**—a component of an analytic technique that enables execution of the technique but does not provide a conclusion or assessment by itself. Analytic tools facilitate the application of techniques by allowing analysts to display or arrange information in a way that enables analysis. An example of an analytic tool is a link diagram or a matrix. Not all analytic techniques have associated tools.
- **Method**—a set of principles and procedures for conducting qualitative analysis.

STRUCTURED ANALYTIC TECHNIQUES

C-2. Structured analysis helps ensure the analytic framework—the foundation upon which analytical judgments are formed—is as solid as possible. It entails separating and organizing the elements of a problem and reviewing the information systematically. Structured analytic techniques provide ways for analysts to separate the information into subsets and assess it until they generate a hypothesis found to be either feasible or untrue (ATP 2-33.4). Structured analytic techniques can drive analytic planning by identifying indicators and evidence that might answer the analytic problem. Cognitive techniques include—

- Basic.
- Diagnostic.
- Advanced.
- Contrarian.
- Imaginative.

C-3. Figure C-1 on page 30 provides a summary of structured analytic techniques. For detailed information about the application of structured analytic techniques, refer to ATP 2-33.4.

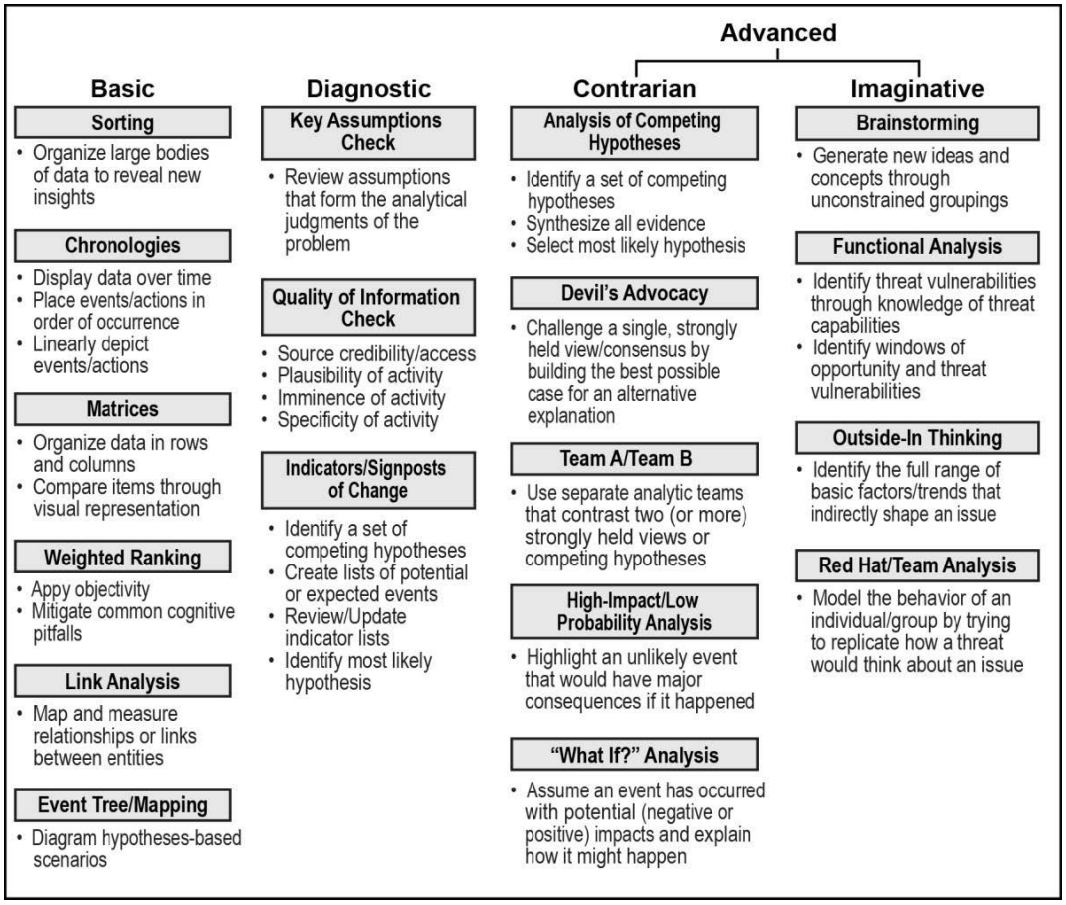

Figure C-1. Summary of structured analytic techniques

ADDITIONAL ANALYTIC TECHNIQUES

C-4. Analysis may also involve the manipulation of data or information in objective manners using mathematical operators. These techniques are primarily from the field of data science and may sometimes be referred to as tools or methods. They are useful for the large datasets common in the cyberspace domain. However, the data size and computations involved make these techniques impractical or infeasible to perform without the use of software tools.

ANOMALY DETECTION

C-5. Anomaly detection is one of the most fundamental analytic techniques. It involves the detection of rare or unusual data or events. Anomalies can be deduced from a set of universal rules or inferred given empirical observations about the environment. For instance, if server maintenance is supposed to occur at 0300 on Tuesdays, software installation at 1800 on Saturday may be an anomaly. Alternatively, if 99 percent of Domain Name System packets are 512 bytes or less, a packet over 1,024 bytes may be an anomaly. Anomaly detection often relies on additional techniques.

DESCRIPTIVE OR SUMMARY STATISTICS

C-6. Descriptive statistics, also called summary statistics, provide general information about data. Common statistics include minimum, maximum, count, measures of centrality (mean, medium, and mode), and measures of spread (range, interquartile range, variance, and standard deviation). While these statistics may yield immediate insights, such as the maximum network traffic coming from an unexpected system, they are generally used in conjunction with other techniques.

OUTLIER DETECTION

C-7. Outlier detection is the process of finding data points that deviate significantly from other points. Through outlier detection, analysts may learn about the operational environment or identify anomalous activity that could indicate threat activity. For instance, outlier detection on central processing unit utilization may reveal cryptocurrency mining. Similarly, web servers are often outliers for outbound network volume. Detecting an outlier is usually based on the distance or loss between a data point and an expected value exceeding a threshold. A simple method is to classify any data point greater than three standard deviations from the mean an outlier. However, this method is less effective when data does not follow a normal Gaussian distribution.

REGRESSION ANALYSIS

C-8. Regression analysis, and more generally supervised learning, allow an analyst to model the relationship between different variables within a data set. Modeling network activity as a function of time may reveal temporal patterns that are interesting or for which the deviations are interesting. For instance, hypertext transfer protocol traffic 1,000 times higher than predicted for a certain time of day may indicate a distributed denial of service or other event relating to that network's mission.

CLUSTERING

C-9. Clustering involves gathering groups' data points based on their attributes. K-means is one method of clustering. These clusters can reveal information about the environment or enable more granular understanding of specific clusters. For instance, clustering users may result in a cluster that contains predominantly network administrators. By looking at the non-administrators in the cluster, analysts may discover administrators that were not previously reported, users with unauthorized privileges, or merely a benign coincidence.

ENRICHMENT

C-10. Enrichment enables greater understanding by adding additional information to a data set. An analyst may enrich network traffic with Domain Name System logs to identify the domains associated. Signature-detection can also be a form of enrichment as the signature list is combined with another dataset to determine which entries match the signatures. Enrichment often uses database joins, lookup tables, and set operations.

This page intentionally left blank.

Source Notes

This division lists sources by page number. Where material appears in a paragraph, it lists the page number followed by the paragraph number.

Page 6, Paragraph 1-25 and 1-26. MITRE ATT&CK®. https://attack.mitre.org/.

Page 9, Paragraph 2-10. Bianco, David J., "*The Pyramid of Pain*", 17 January 2014, http://detect-respond.blogspot.com/2013/03/the-pyramid-of-pain.html.

Page 9, Figure 2-3. Bianco, David J., "*The Pyramid of Pain*", 17 January 2014, http://detect-respond.blogspot.com/2013/03/the-pyramid-of-pain.htmlPage 13, Paragraph 3-15. MITRE ATT&CK®. https://attack.mitre.org/.

Page 15, Paragaraph 4-5 and 4-6. MITRE ATT&CK®. https://attack.mitre.org/.

Page 25, Paragraph B-1 and B-2. MITRE ATT&CK®. https://attack.mitre.org/.

Page 25 Note. MITRE ATT&CK®. https://attack.mitre.org/.

Page 26, Paragraph B-5 and B-7. MITRE ATT&CK®. https://attack.mitre.org/.

Page 27, Paragraph B-8. MITRE ATT&CK®. https://attack.mitre.org/.

This page intentionally left blank.

Glossary

The glossary lists acronyms and terms with Army or joint definitions. The proponent publication for terms is listed in parentheses after the definition.

SECTION I – ACRONYMS AND ABBREVIATIONS

ADP	Army doctrine publication
ATP	Army techniques publication
ATT&CK®	adversarial tactics, techniques, and common knowledge
CJCSM	Chairman of the Joint Chiefs of Staff manual
DA Form	Department of the Army form
FM	field manual
JP	joint publication
NIST SP	National Institute of Standards and Technology Special Publication

SECTION II – TERMS

analysis

The compilation, filtering, and detailed evaluation of information to develop knowledge or conclusions. (ATP 2-33.4)

analytics

The systemic processing and manipulation of data to uncover patterns, relationships between data, historical trends, and attempts at predictions of future behaviors and events. (NIST SP 1500-1)

measure of effectiveness

An indicator used to measure a current system state, with change indicated by comparing multiple observations over time. (JP 5-0)

measure of performance

An indicator used to measure a friendly action that is tied to measuring task accomplishment. (JP 5-0)

This page intentionally left blank.

References

All URLs accessed on 10 July 2023.

REQUIRED PUBLICATIONS

These documents must be available to intended users of this publication.

DOD Dictionary of Military and Associated Terms. September 2023.

FM 1-02.1. *Operational Terms*. 9 March 2021.

FM 1-02.2. *Military Symbols*. 18 May 2022.

RELATED PUBLICATIONS

These documents contain relevant supplemental information.

JOINT PUBLICATIONS

Most joint publications are available online: https://www.jcs.mil/doctrine.

JP 5-0. *Joint Planning*. 1 December 2020.

ARMY PUBLICATIONS

Most Army doctrinal publications are available online: https://armypubs.army.mil.

ADP 6-22. *Army Leadership and the Profession*. 31 July 2019.

ATP 2-33.4. *Intelligence Analysis*. 10 January 2020.

FM 3-55. *Information Collection*. 3 May 2013.

FM 6-27/MCTP 11-10C. *The Commander's Handbook on the Law of Land Warfare*. 7 August 2019.

OTHER PUBLICATIONS

National Institute of Standards and Technology special publications are available online: https://csrc.nist.gov/publications.

NIST SP 1500-4. *Big Data Interoperability Framework*, 21 October 2019.

RECOMMENDED READINGS

FM 3-12. *Cyberspace Operations and Electromagnetic Warfare*. 24 August 2021.

WEBSITES

This section contains no entries.

PRESCRIBED FORMS

This section contains no entries.

REFERENCED FORMS

Unless otherwise indicated, DA Forms are available on the Army Publishing Directorate Website: https://armypubs.army.mil.

DA Form 2028. Recommended Changes to Publications and Blank Forms.

This page intentionally left blank.

Index

Entries are by paragraph number.

This page intentionally left blank.

TC 3-12.2.4.1
17 JANUARY 2024

By Order of the Secretary of the Army:

RANDY A. GEORGE
General, United States Army
Chief of Staff

Official:

MARK F. AVERILL
Administrative Assistant
to the Secretary of the Army
2400907

DISTRIBUTION:
Active Army, Army National Guard, and United States Army Reserve. Distributed in electronic media only(EMO).

This page intentionally left blank.

PIN: 216865-000