NPS-AM-21-034

ACQUISITION RESEARCH PROGRAM SPONSORED REPORT SERIES

Acquisition Data Analytics for Supply Chain Cybersecurity

February 26, 2021

Dr. Randy Maule, Research Associate Professor

Graduate School of Operational & Information Sciences

Naval Postgraduate School

Approved for public release; distribution is unlimited.
Prepared for the Naval Postgraduate School, Monterey, CA 93943.

ACQUISITION RESEARCH PROGRAM
GRADUATE SCHOOL OF DEFENSE MANAGEMENT
NAVAL POSTGRADUATE SCHOOL

The research presented in this report was supported by the Acquisition Research Program of the Graduate School of Defense Management at the Naval Postgraduate School.

To request defense acquisition research, to become a research sponsor, or to print additional copies of reports, please contact the Acquisition Research Program (ARP) via email, arp@nps.edu or at 831-656-3793.

ACQUISITION RESEARCH PROGRAM
GRADUATE SCHOOL OF DEFENSE MANAGEMENT
NAVAL POSTGRADUATE SCHOOL

Abstract

Cybersecurity is a national priority, but the analysis required for acquisition personnel to objectively assess the integrity of the supply chain for cyber compromise is highly complex. This paper presents a process for supply chain data analytics for acquisition decision-makers, addressing data collection, assessment, and reporting. The method includes workflows from initial purchase request through vendor selection and maintenance to audits across the life cycle of an asset. Artificial intelligence can help acquisition decision-makers automate the complexity of supply chain information assurance.

Research Objective: This research discusses techniques, systems, tools, and workflows to empower acquisition departments for supply chain data analytics for cybersecurity compliance and information assurance.

Research Questions: Can tools for supply chain data analytics be integrated into acquisition workflows to address information assurance across the supply chain? Will information assurance models be sufficient for acquisition departments to implement cyber controls in the purchase process, across the supply chain, and throughout the life cycle of an asset?

THIS PAGE LEFT INTENTIONALLY BLANK

About the Author

Dr. Randy Maule has been with the Naval Postgraduate School since 2000, serving as an enterprise developer and technical analyst in joint forces and coalition exercises where he managed knowledge systems and conducted test and measurement. His enterprise tool suite and technical analysis systems operated on ships, in maritime and network operations centers, and in forward-deployed commands for nearly 15 years. Prior to this, he spent 10 years in Silicon Valley working with intelligent networks and service architecture in the telecommunications industry, and prior to this, developing enterprise knowledge systems at a federal supercomputer center.

THIS PAGE LEFT INTENTIONALLY BLANK

NPS-AM-21-034

ACQUISITION RESEARCH PROGRAM SPONSORED REPORT SERIES

Acquisition Data Analytics for Supply Chain Cybersecurity

February 26, 2021

Dr. Randy Maule, Research Associate Professor

Graduate School of Operational & Information Sciences

Naval Postgraduate School

THIS PAGE LEFT INTENTIONALLY BLANK

Table of Contents

THIS PAGE LEFT INTENTIONALLY BLANK

Introduction

The analysis required for acquisition personnel to objectively assess the integrity of the supply chain for cybersecurity risk is complex and multifaceted, requiring many different types of expertise. Audit duties have not traditionally been the purview of acquisition departments. Yet, as the arbiter in purchase decisions across the life cycle of an asset, we may logically extend the acquisition role to accommodate information assurance audits—especially given recent failures from stovepipe approaches that have left information assurance gaps.

The common element across the life cycle of a technology—from initial requisition, through deployment and implementation, to contracting for technical support and maintenance—is the acquisition role. However, this role does not generally have the resources or technical expertise to accomplish comprehensive and continuous cybersecurity assessment across the supply chain and life cycle of an asset. New analytic methods coupled with artificial intelligence (AI) automation support such an extension.

In this paper, we extend a method for supply chain data analytics successfully deployed in the Fleet for over a decade to the acquisition role. The method includes workflows for acquisition decision support from initial purchase request through vendor selection and maintenance across the life cycle of an asset. The research includes a discussion of integrity analysis, data analytic tooling and reporting methods, and AI automation to help acquisition decision-makers manage the complexity of supply chain cyber analysis for information assurance.

THIS PAGE LEFT INTENTIONALLY BLANK

Background

Naval and joint forces systems analysis has identified significant risks within the supply chain (Military & Aerospace Electronics, 2020). Compromised chips, vendor part substitution, and post-purchase systems maintenance all contribute to the problem (Villasenor & Tehranipoor, 2013).

Various techniques have been developed to address the problem, ranging from the National Institute of Standards and Technology (NIST) Risk Management Framework (RMF; National Institute of Standards and Technology, 2020a) and Cyber Security Framework (CSF; National Institute of Standards and Technology, 2020b), to the Cybersecurity Maturity Model Certification (CMMC) process (Carnegie Mellon University and The Johns Hopkins University Applied Physics Laboratory, 2020), to the Department of Defense (DoD) Digital Modernization Strategy and its Development (Dev) Security (Sec) Operations or DevSecOps principles (Department of Defense, 2019), to the Fleet's Compile to Combat in 24 hours (C2C24) process for secure software (SPAWAR Office of the Chief Engineer, 2018).

This paper provides a means to operationalize these frameworks, standards, and principles and extend the analysis to include supply chain assessment across the life cycle of an asset. Our intent is to extend the method used for over a decade for technology evaluation in the FORCENet Sea Trials into the acquisition role to support decision-makers. We successfully employed the methodology to decompose hardware and software specifications for cybersecurity and supply chain analysis, execute test and measurement on the FORCENet technologies in various operational contexts, and conduct systems integration analysis during the Sea Trials and supporting joint forces and coalition exercises.

In our previous research, we presented the FORCENet analytics framework and extended that capability to the acquisition role for supply chain management across the life cycle of an asset (Maule, 2019a). With budget authority, the acquisition role is logically positioned to oversee the supply chain—from initial requisition to installation and maintenance. The complexity of the task and the multiple levels of technical

expertise required can be mitigated through software and automation. This project extends the previous research with AI algorithms to enhance the automation.

AI integrated into acquisition decision support systems provides a solution (Maule, 2020). This involves automation of the test and measurement process, together with expert systems for equipment, vendor, and contract selection. We then render the results into a dashboard for easy understanding by acquisition decision-makers.

Acquisition

We previously established the variables and metrics for supply chain cybersecurity assessment from purchase request, through vendor selection, to maintenance audits (Maule, 2019b). We deployed this as an “analytics grid” (Maule, 2014) across ships and shore commands for FORCENet systems assessment for over a decade (Maule & Lewis, 2011). The next step is to extend these capabilities to the acquisition role where supply chain information assurance can be addressed from a budget perspective. Our method advances multi-disciplinary research to evaluate all variables that we found to impact the validity of naval systems and data under operational load and within technical and environmental context (Maule, 2016).

A difficulty is that the relationship between systems, components, and other systems is typically nonlinear. It is not possible to precisely define the inputs such that there is a direct relationship to the outputs. Cause–effect relationships are probabilistic and can be determined only within technical, operational, and environmental context (Maule, 2015). Additionally, systems performance tends to exhibit divergent patterns under stress—such as challenged communications, jamming or electronic attack, and of course cyber and electronic manipulation. Similar to our naval exercises, the acquisition decision support software will need to address complex operational scenarios.

Adaptive systems are characterized by the capability to learn from experience. Machine and deep learning are an examples of adaptive systems, and these tools can be applied to help understand complex relationships. We observe adaptive behaviors in naval exercises as we instrument networks to monitor complex data flows across geographic regions. Components of systems interact, with the result of those

interactions dependent on dynamic variables, for example, changes made as services adapt to new tactical scenarios.

Our evaluation addresses the dynamic interplay of variables, in context, over time. Failure to address this complexity results in an inability to recognize performance variance or cyber compromise, and therein to adapt the analysis to changes in operational or environmental context. So, the assessment is not a static point in time for a single workflow but across the life cycle of an asset including an evolving supply chain. This is what differentiates our methodology from others that tend to evaluate over a defined period of time and not across the supply chain and life cycle of the asset in different technical, operational, and environmental contexts.

In the following sections, we discuss digital tooling and analytic methods for supply chain integrity analysis that includes data collection, results visualization, and decision support. The acquisition role is advanced as an arbiter of supply chain viability. AI tools and algorithms integrated into the acquisition workflows help acquisition decision-makers address cybersecurity compliance along with information assurance across the supply chain over the life cycle of an asset.

Decision Support

The level of decision support required for a task as complex as cybersecurity assessment across the supply chain and throughout the life cycle of an asset is realized through AI and automation. Probabilistic algorithms have been found to be effective in addressing multiple dimensions of analysis when contexts are dynamic and expanding (McMullen, 2015), such as naval operations.

In Figure 1 we summarize major trends that have evolved AI capabilities, beginning with the early works in general intelligence for computer decision support, to content (Goldberg et al., 1992) and group filtering (Resnick et al., 1994), to the recommender systems (Resnick & Varian, 1997) popular in today's online search and commerce (Koren, 2008).

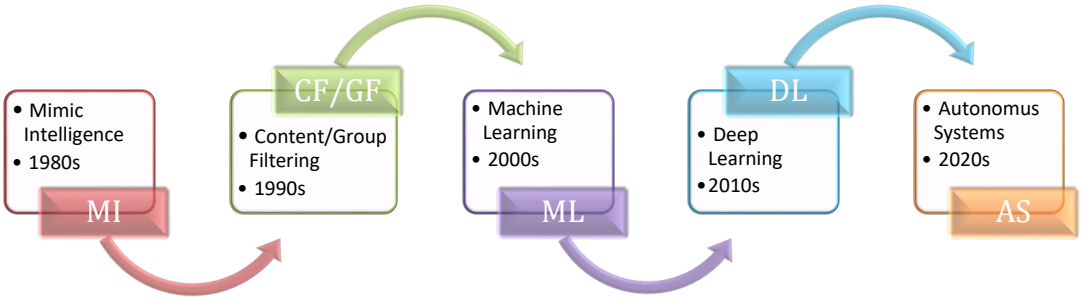

Figure 1. Evolution of AI for acquisition decision support

A DoD example is the content filtering AI that we used on a daily basis for technology assessment in the FORCENet exercises, integrating quantitative network and system technical measures with qualitative user analysis to determine the status of an asset in context (Maule et al., 2010). An extension of this AI to include machine and deep learning with autonomous collection and processing will support the proposed acquisition supply chain decision support system. The more AI in the workflow, the greater the automation. For their part, decision-makers will review portals that aggregate the results, visualizing data and monitoring alerts and notifications.

Machine learning algorithms add user behavior into the analysis, building capabilities by example rather than direct programming (Dietterich, 2003). These algorithms excel in settings where specifications for desired program behavior are not available but where examples of this behavior are available.

Figure 2 illustrates the different programming approaches. Expert systems, filtering, and decision support were a facet of early AI but are still popular in today's search engines and recommendation algorithms. Machine and deep learning have history back to the 1980s, but recent applications have popularized these algorithms for a variety of tasks, including in our use case for structured reinforcement of self-learning in supply chain decision support.

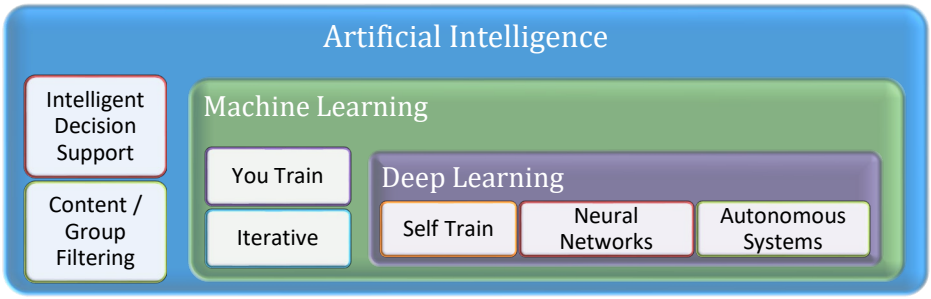

Figure 2. Machine and deep learning for acquisition decision support

In machine learning, the algorithm is taught to make an accurate prediction. An operator provides guidance through positive and negative reinforcement—for example, to determine if an anomalous behavior has occurred or a data synthesis is incorrect. Deep learning is a compute-intensive subset of machine learning where the algorithm learns to predict by itself using a brain-like artificial neural network structure (Farsal et al., 2018). The algorithms extract features independent of an operator (Microsoft, 2020a). In machine learning, small amounts of data can train the machine to make predictions so machine learning is more appropriate to begin this work. In recent tests, we achieved over 85% prediction accuracy with as few as a dozen data sets.

Figure 3 provides an example of the machine and deep learning workflow, beginning with data collection and processing, applying the algorithm, then evaluating the results. In the machine learning process, the data is split in the training phase, generally about a 75/25 split—with the former to train the model that is applied to the latter. The process iterates until the correct decisions are achieved, at which point the algorithm is implemented.

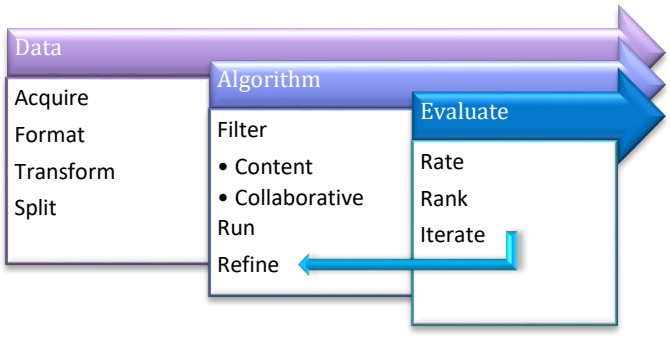

Figure 3. Decision support machine learning workflow

An additional benefit for acquisition decision support is that machine learning can model the necessary expertise from the multiple technical specializations required for comprehensive supply chain analysis. Models are trained by experts and then integrated into the decision support system. With the addition of deep learning, the models can independently evolve, continuously learning to improve decision support with ever more efficient predictions. The acquisition decision-maker is viewing the high-level report or status visualization in the portal, looking for any alerts or notifications that have been automatically generated from the analysis that may indicate an abnormality or deviation from expected patterns.

Method

We begin the workflow with conceptual models and architecture and then add procedures for technical assessment with tooling for data monitoring and analytics. We implement applications for in-service audits for cybersecurity and information assurance, systems verification, and data validation. Note that few of the prevailing cyber methodologies take this final step of validating the data against baselines, user actions, and in-service comments from operators. Finally, the technical models and their measurements are integrated with audit workflows for comprehensive life-cycle systems assessment. In the FORCENet, we employed measurements for baseline operations, under load, under stress, and under electronic attack.

We implemented algorithms to compile the generated data and associated knowledge into the portals for decision-makers. Collectively these techniques help acquisition decision-makers from (a) initial purchase request, to (b) vendor selection, through (c) implementation and operations, to (d) long-term maintenance and contractor evaluation. AI automates the analysis and manages the alerts and notifications that are sent to decision-makers.

THIS PAGE LEFT INTENTIONALLY BLANK

Analytics

Previous research established a workflow for acquisition supply chain audits (Maule, 2019b). On the left side of Figure 4 we illustrate this workflow, showing the pre-acquisition process before and after cybersecurity enhancement (box “A”). Box “B” shows the addition of test and measurement (T&M) within the workflow for monitoring, maintenance, and compliance reporting across the life cycle of the asset. All are rendered into dashboards for decision-makers.

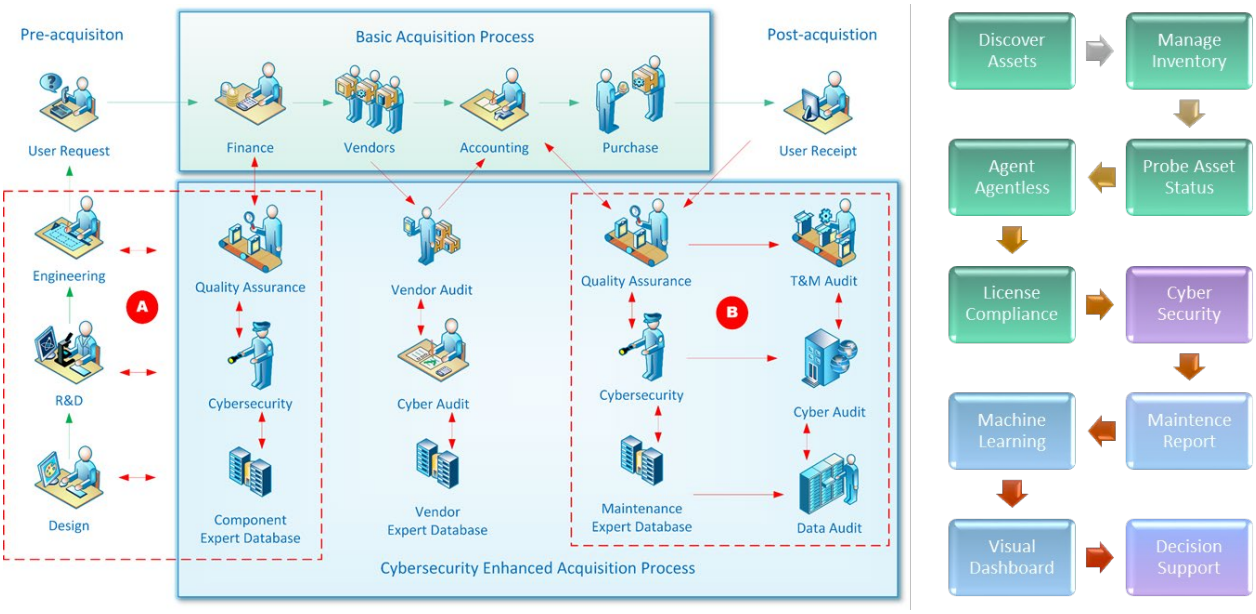

Figure 4. Acquisition supply chain cybersecurity test and measurement workflow

On the right side of Figure 4, we show the operationalization of the workflow that is the focus of this paper. Specifically, the process through which assets are discovered and analyzed for compliance. The workflow includes asset management, management of test processes, generation of results, and automation of the manual aspects of the decision support process as appropriate. Acquisition decision-makers are provided with notifications and alerts in the portal with links to the tools for more detailed analysis if needed.

Architecture

We implement the decision support portal to aggregate reports from the monitor and measurement tools. Enterprise analysis typically begins with the collection of system logs (Syslogs) to help determine the status of each system. The collection and analysis are automated and the results rendered to the decision portal. After an incident has occurred, the Syslogs are compiled with other data to achieve a detailed understanding of the problem within its operational context. Syslogs are correlated with communication logs from switches, routers, firewalls, and gateways for both radio frequency (RF) and wired connections. The human and contextual elements are then factored into the analysis. Events are correlated against the analysis models to provide integrated monitoring and prediction. This process is automated so only high-level visualizations and reports reach the acquisition decision-makers. If more detailed information is required, then events can be queried for specifics that can be forwarded to specialists for forensic analysis.

Tools with this capability include application performance monitors (APM), information technology (IT) infrastructure monitors (ITIM), network performance monitors with diagnostics, and digital experience monitors (Rich et al., 2019). Figure 5 provides an example of tools integrated into the architecture for data collection, analysis, and decision support. This is the architecture in operation under the decision support portal.

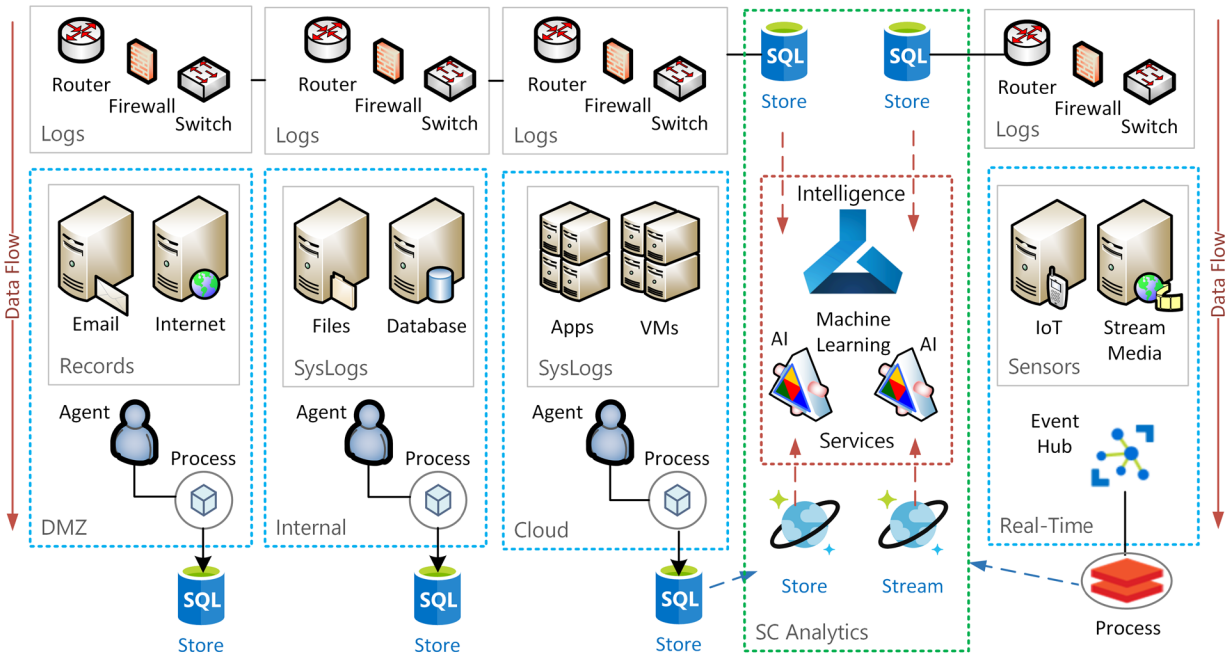

Figure 5. Acquisition supply chain assessment architecture

Communication analysis begins with data drawn from routers, firewalls, switches, gateways, and interconnection devices. Network tools aggregate metadata into storage and then transfer that data to analytic engines where events are correlated and results visualized. The better tools have multiple levels of built-in AI to evaluate patterns over time, automatically detect anomalies, and send alerts or notifications to decision-makers. At this level, the acquisition decision-makers are only concerned with the reports that are automatically linked to the portal and specifically the alerts and notifications that automatically populate to the portal.

There are both agent and agent-less approaches to collect data. Both are automated, but the systems onto which intelligent software agents are loaded tend to produce more insightful information. The more comprehensive analytic tools provide a tight linkage between report consoles and collection agents, with the agent not only collecting data but processing that data before storage or routing to the decision support dashboard(s).

Intelligent agents (software) are best for systems with access to the public Internet, for example, in a demilitarized zone (DMZ) within a company. This helps

protect internal systems while simultaneously detecting and recording cyber incidents. Streaming services can use a network “tap” or signal duplicator and an agent-less approach to not interfere with time-critical data where milliseconds can make a difference. Cloud operations tend to be between public and private, depending on the configuration and security controls.

On the far right of Figure 5 are the real-time and streaming services, typically from sensors, video, audio, news feeds, and other just-in-time media. Internet of Things (IoT) and edge devices such as cell phones and other mobile devices would be included. The intent is real-time analysis, so data is streamed through an event hub to a processing node configured with real-time analytics. Data is simultaneously streamed into the acquisition supply chain analytics workflow and decision support dashboards.

Collection

All components and facets of the supply chain are monitored and assessed, including personnel. Test and measurement (T&M) begins with asset discovery from sensors that are collecting information about technical, operational, and environmental factors that will impact supply chain analysis. For reference, we assume a tactical application with live, deployed assets.

In Figure 6 we present the radio and network sensor monitors that provide the first step to the contextual understanding necessary for our analysis. Sensors monitor the network devices in the architecture. Once again, decision-makers are not involved at this level, only with the alerts and notifications that may arise from this level, and only then to be aware that a system is no longer viable and will need to be replaced or maintenance personnel contracted. The decision-maker will need to be cognizant of where the equipment is in its life cycle to balance the cost of continued maintenance versus the price of new equipment.

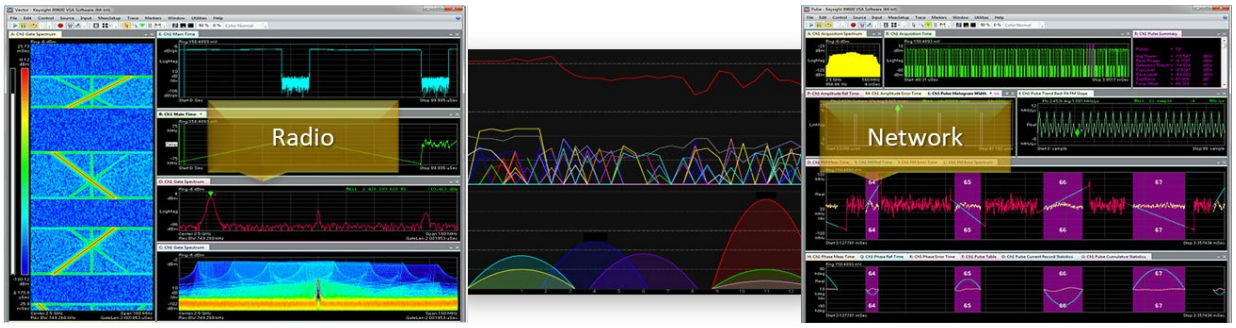

Figure 6. Spectrum radio and network data monitor and capture

Sensors can be grouped by region and reported by mission. Most of the better measurement tools have this capability. RF and network collection devices monitor and report on asset performance, interference, and operational issues such as cyber and electronic attack. T&M tools automatically discover and classify assets, determine data and network conditions, and analyze data usage patterns and users for potential risk. Again, this is not a concern for acquisition decision-makers, but the reports from these devices should factor into purchase decisions as they may indicate a current need or life-cycle element.

We train AI algorithms to observe normal patterns and once trained to spot abnormal behaviors, and provide alerts and notifications. The more capable tools offer predictions that may help decision-makers. An example is the maintenance cycle and the mean time to failure—at which point a major repair or replacement is needed, so the budget needs to have a provision for recurring expenses. Intelligent agents aggregate metrics across data streams and perform both real-time assessment and cognitive/predictive analysis. Acquisition decision-makers review the reports when purchase requests are initiated, or when alerts and notifications indicate a need for assessment.

After being pulled from the radio, network sensors, and systems, data are dissected and analyzed (Figure 7). Our analysis includes control logs, user roles and services, and security policies. This process is automated so acquisition decision-makers merely access reports as needed. Applications are monitored through passive and active taps, agents, and Syslogs. Asset monitors validate license compliance for

information assurance. Service monitors assess processing routes, message transformations, and data integration.

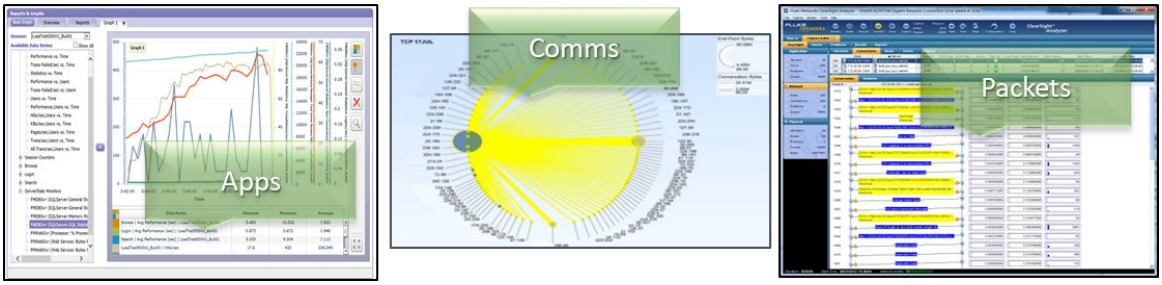

Figure 7. Sensor communication, application, and packet analysis

Packets are decomposed to analyze their data. Status reports, along with real-time event streams, can be aggregated for continuous assessment. In addition to status reports, the acquisition decision-maker can apply predictive algorithms to determine whether systems or data have been compromised. The center image in Figure 8 is an example of a high-level report visualization that would be presented in the portal for the acquisition decision-makers. If the decision-maker noted an anomaly and desired more information, they would access the more detailed reports.

Once events and data are normalized and correlated, the analytics detect data and user anomalies (Figure 8). Cognitive algorithms assess usage patterns to derive behavior profiles, integrating user, system, and network information to provide decision intelligence. This too is somewhat unique to the FORCENet methodology wherein we integrated warfighter chat logs with technical performance measures to detect problems even before the equipment issued alerts.

Figure 8. Event analysis through performance and function monitors

Automated agents at each of the tactical nodes aggregate and filter data, then synchronize with the acquisition supply chain analytic servers. Data science is automatic as resources pass through the tactical processing nodes. This includes cyber status along with cyber test and measurement (Figure 9).

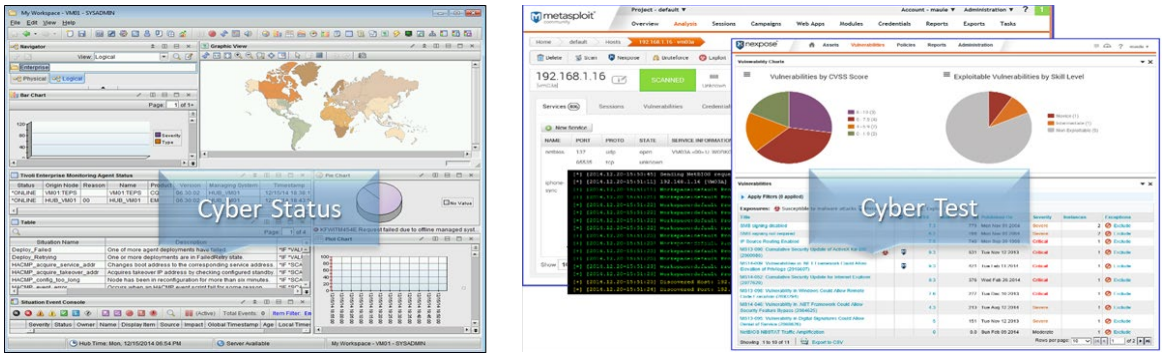

Figure 9. Cyber status monitors with automated cyber stress tests

Analysis

Acquisition decision-makers can review the quantitative network, system, user, and transaction dashboards. In a typical day, they would just monitor the dashboards for any automated notifications or alerts that may indicate a deviation has been observed. Predictive analytics, also automated, correlate events and context over time to quantify the status of the monitored event within its operational context. Tests for each component within a system are evaluated for internal and external operations, compiling the data from the tools mentioned above into a score.

This is the manner in which we conducted Fleet tests within technical, operational, and environmental context—for example, independent when pier-side without communications, again under full communications load, and again in different

operational scenarios. A system, its code, and its interfaces to sensors and networks may perform perfectly fine in static tests, but the performance will be much different in an operational context when subjected to degraded, intermittent, and limited (DIL) communications.

Such conditions may be experienced, for example, if those systems are active in anti-access/area denial (A2AD) electronic exploits or are under cyberattack. Unfortunately, it is often not until these taxing operational scenarios that we determine there is a problem. The goal of the research herein, and the expanded acquisition decision support role, is to preempt such occurrences.

We successfully used the above process in FORCENet systems evaluations for over a decade. However, the methodology and techniques were never extended to help acquisition personnel make the supply chain decisions that could have preempted many of the problems. The escalation of nation-state cyber activity and the public awareness of supply chain vulnerabilities necessitates such an extension to the acquisition role and a rigorous approach to the acquisition analysis process.

Finally, we compile quantitative results into a decision support matrix (Figure 10). A coefficient is derived from test results, and based on the number of tests in a given context, a confidence level is rendered for that system, in that context, at that point in time. This provides the basis for the installation and maintenance audit phases of the acquisition supply chain cybersecurity T&M workflow.

Lifecycle Phase Analysis: System Readiness by Technical, Operational, Environmental Confidence

	Code	System	Sensor	LAN	Spectrum	WAN	NOC	GIG	Coefficient	Tests	Confidence
Laboratory	0.85	0.85	0	0.9	0	0	0	0	0.325	12	4%
Pier-side	0.85	0.85	0	0.88	0	0.8	0.85	0.83	0.6325	22	14%
Scenario 1: Load	0.8	0.8	0.76	0.87	0.73	0.75	0.71	0.72	0.7588	8	60%
Scenario 2: DIL	0.77	0.67	0.58	0.88	0.6	0.51	0.58	0.52	0.63875	47	30%
Scenario 3: A2AD	0.6	0.58	0.49	0.85	0.28	0.37	0.25	0.28	0.4625	53	25%
Scenario 4: Cyber	0.24	0.33	0.49	0.6	0.28	0.37	0.25	0.28	0.355	31	11%

Analysis

Readiness Assessment Coefficient = (Lifecycle Area Measures) / (Measured Areas)
Readiness Confidence = (Lifecycle Area Measures) / (Measured Areas) * (Independent Context Evaluations)
Perfect Readiness Assessment Confidence = (8) / (8) * (100) = 100%

Figure 10. Decision support matrix with confidence levels

Automation

This section helps us monitor hardware, software, and service dependencies in real time. The intent is to implement algorithms that dynamically learn about the data and then correlate that data to assess patterns and spot anomalies from those patterns (BMC, 2020). Once programmed, the monitors will predictively alert and help remediate issues before they impact the services and budget.

In process, data sources are aggregated and analyzed using machine learning algorithms trained on the data. Then, the trained model is deployed (Figure 11).

Figure 11. Audit automation sequence

After deployment, the analysis continues automatically until a revision is necessary. Alerts or notifications from audit assessment results are populated to a dashboard for supply chain decision-makers.

Algorithms

The previous sections have discussed the use of AI and automation but not the specific algorithms that are key to both. We cover two basic categories of AI in the following sections—supervised (machine learning) and unsupervised (deep learning). A human operator is active in the supervised algorithm when the data is categorized and labeled. The algorithm studies the labeled examples and then makes recommendations or predictions. In unsupervised or deep learning, there are no labels. Instead the data is organized, and the algorithm learns by itself. This requires large data sets, computing power, and time.

Regression algorithms are usually the first step in the discovery process and help us determine the suitability of the data sets. Simple linear regression is used to model the dependent and independent variables to determine whether a linear function is viable for our use case, for example, to determine whether we can predict the

dependent variable as a function of the independent variable (Altman & Krzywinski, 2015). Regression algorithms provide a relatively straightforward means to generate predictions but may lack the flexibility necessary for interpreting complex supply chains with multiple, dynamic variables.

A little more complex but a more robust approach is multivariate linear regression where there can be multiple, correlated dependent variables. Still, this may be insufficient for complex supply chain analysis, such as when a new technology is added into a system-of-systems architecture, a vendor in the supply chain has substituted components, or a maintenance technician has skipped critical steps in installation or repair.

A variant we used in the FORCENet analysis was cube regression where a multi-dimensional array (hypercube) provided a three-dimensional level of analysis, with each dimension containing categorical data.

Another of the basic machine learning algorithms is logistic regression to predict the probability of a certain event. This can be used in image analysis to extract objects and match events with context, for example, the operations of a system within certain operational, technical, or environmental contexts.

Finally, are the deep learning neural nets that support multilayered analysis. These are self-trained algorithms (vice the human-supervised algorithms of machine learning) that can significantly enhance accuracy but require very large data sets, robust hardware, and long training times. At this stage in our supply chain analysis, the deep learning algorithms are probably not appropriate, but we discuss them for future reference when over time those large data sets do become available.

Table 1 provides a high-level summary. Each approach and algorithm has its strength and weakness. Multiple algorithms may link into an analysis workflow. Some algorithms excel at prediction, others at grouping, and others at classification.

Table 1. Algorithm types and function for acquisition decision support

Type	Data Source	Prediction Algorithm	Method
Content Filtering	➢ User ➢ System ➢ Network	Knowledge Graph: ➢ User comments ➢ System actions ➢ Network capture	➢ Performance ➢ Rankings
Collaborative Filtering	➢ Preferences ➢ Associations ➢ Patterns	Memory-based: • Similarity matrix Model-based: • Cluster correlation	➢ Weighted sum ➢ Regression
Deep Learning	➢ Operations ➢ Context ➢ Items	Neural Filtering: • Matrix factorization	➢ Models ➢ Features

Regression algorithms help predict a value, while classification algorithms help predict classes or clusters. Examples of the former include boosted decision trees, decision forests, and linear and neural network regression. Examples of the latter include multiclass decision trees, multiclass decision forests, multiclass logistic regression, and multiclass neural networks (Microsoft, 2020b).

In the proposed supply chain analytics and decision support system, these various algorithms will look for particular events or suspicious activities—such as system anomalies or component failures, irregular data patterns or cyber intrusions, maintenance alerts, etc. Table 2 categorizes some of the basic forms of incident identification and the algorithms optimized to spot these anomalies:

Table 2. Acquisition supply chain anomaly alerts and algorithms

Incident	Algorithms
System alarms	Clustering and pattern matching
Cause analysis	Decision trees, random forest, graph analysis
Abnormalities	Statistical, probabilistic: univariate and multivariate, correlation, clustering, classifying, and extrapolation
Irregular patterns	Correlation/prediction from historical and/or streaming data through clustering or grouping
Context	Topology patterns displayed through graph data to establish relevance and hidden dependencies

Information assurance begins with contextual analysis of the data and visualization of the service environments. Alarms can be managed through clustering and pattern-matching algorithms; root cause through decision trees and graph analysis.

Most professional collection and monitoring tools have built-in AI for abnormality detection with a drill-down into events of interest (BMC, 2018a). So, all we need is a means to aggregate the results to a decision dashboard and integrate those results into our supply chain decision process. Once this is done, the decision support system can automatically predict from correlation. Additionally, many tools provide topology analysis with context to help decision-makers assess linkages and dependencies across applications, networks, and databases.

Data flows through the networks can be monitored for protocols, devices, and users. Details on load, latency, errors, and sessions can be derived. Suspicious events can be correlated with Syslogs and user data (chat, service requests, forum posts, etc.). Results can be rendered into service dashboards for acquisition decision-makers to assess the need for proactive maintenance or replacement. System health status, alarm history, and usage metrics support decisions.

Service dependency maps ensure client–server relationships and system integration contingencies are understood (NETSCOUT, 2019). Compliance can be attested to ensure that the proper number of licenses are active. Power consumption can be integrated for cost–benefit analysis. Decision-makers can generate inventory reports for tracking, implementation, and maintenance.

Once integrated with usage data, then life-cycle metrics can be applied. Device security can be monitored continuously, including mobile devices. If cyber compliance is not in order, the devices can be locked or wiped (BMC, 2018b).

Since intelligent analytics, machine learning, and other probabilistic techniques are prone to false positives, bias, and other errors, the supply chain analytics process will continuously test for efficiency—hence the need for automation. Over time the machine and deep learning algorithms will improve to enable further automation—even to the point of being able to correlate seemingly non-connected events (Ixia, 2017).

Use-Cases

Some algorithms are optimized for specific use-cases within the supply chain acquisition workflow. For example, algorithms can generate recommendations based on user transaction history. An affinity matrix captures the strength of the relationship, considers information about the type of user-item interaction through differential weighting, and ties this to time information about when a user-item event occurred (Microsoft, 2020d). In operation, the algorithm evaluates user–item affinity with item–item similarity to render recommendations.

Alternating Least Squares (ALS) is a collaborative filtering (matrix factorization) algorithm. When coded in Python/Spark, this machine learning algorithm optimizes large-scale distributed data analysis (Apache Spark, 2020) to provide user context for technical evaluation, for example, querying logs from support forums, collaboration sessions, or chat logs.

The R programming language is popular for statistical computing and data analysis (R, 2020). R supports matrix arithmetic for vectors, matrices, arrays, data frames, and lists (Dalgaard, 2002). R's extensible object system includes objects for regression models, time-series and geo-spatial coordinates. In the supply chain acquisition system, the applications may include user polls and surveys, technical specification reviews, etc.

Fraud Detection

A particularly acute problem in the DoD purchase process is fraudulent vendors, substituting used, counterfeit, or compromised products or components into the asset. Fraud detection algorithms are some of the earliest used for machine learning, yet are not generally integrated into the naval acquisition process. This is typically a binary classification problem but in a complex supply chain across the life cycle of the investment, the identification becomes multidimensional:

- Is this OEM equipment? (component database)
- Is this vendor legitimate? (vendor database)

- What is the supply chain for this part/contract? (tracking database)

Sentiment Analysis

In a workflow for fraud detection, the above may be supplemented with sentiment analysis algorithms to assess user reports about the equipment (chat, email, support forums, etc.). This algorithm was used extensively in the FORCENet Sea Trials to help assess sailor sentiment about the new technologies we were evaluating. Problems would often become noticeable in the chat logs long before we had time to review any Syslog alerts. In process, the key variables are extracted, and then a model is trained to look for positive or negative sentiments (Figure 12).

Figure 12. FORCENet sentiment analysis dashboard

Vector Search

Stemming and fuzzy logic can be helpful to summarize documents and even large media repositories (Oracle, 2020). We used a “gist” algorithm to characterize FORCENet technologies in a large knowledge management system with a “theme” algorithm to extract capabilities. In the supply chain acquisition system, these tools can be used to search across databases for relevant information about a technology purchase or its use (Figure 13). Approximate nearest neighbor algorithms build on the generated vector indices to enable a very rapid retrieval of the search data (Microsoft, 2020e).

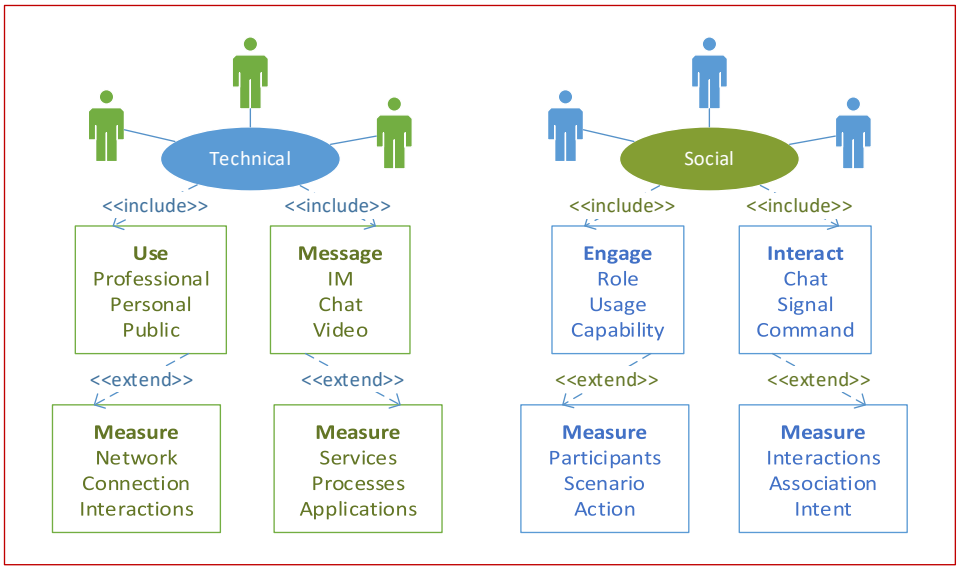

Figure 13. Supply chain assessment variables

Integration of sentiment analysis and cognitive search enable finely filtered data discovery with rich multimedia visualization (Unify, 2020). Cognitive algorithms assess data patterns and derive behavioral analytics, integrating user social media (chats, forums, wikis, etc.) with system and network information to provide decision intelligence. Predictive algorithms spot patterns for decision support about searched technologies.

Recommendation

A workflow for the acquisition supply chain recommendation system is presented in Figure 14. Data is split by time or a selected variable for machine learning, then routed into one or more of the previously discussed algorithms. A human evaluator then judges the recommendation for accuracy and intent. If it is viable, then it is deployed.

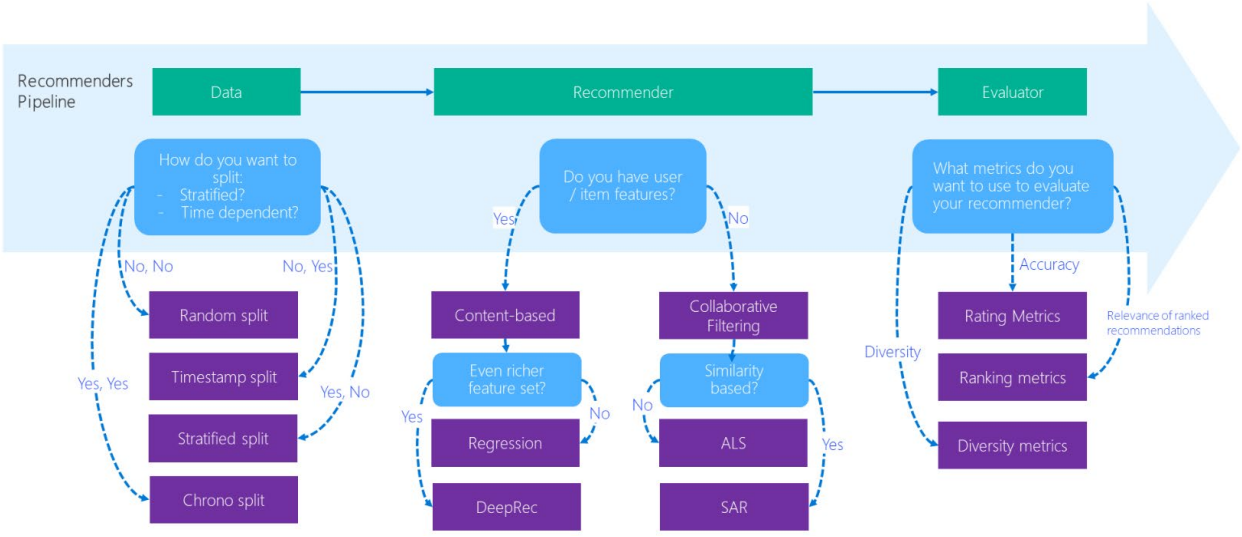

Figure 14. Recommender decision support workflow. Source: Microsoft, 2020a

Integration

We now apply these algorithms to the supply chain problem to provide insight for acquisition decisions. The algorithms may span multiple variables in the supply chain acquisition decision support system. The basic categories include

- Item/text/image/pattern recognition
- Anomaly detection and alert
- User sentiment analysis.

These general classes of capabilities can then be applied to specific supply chain assessment problems to address

- Equipment failure
- Human/systems interfaces
- Viable/not viable for operations.

Table 3 summarizes the major algorithms in the supply chain acquisition system (Microsoft, 2020c). Most are derived from recommendation systems to assist in product selection, installation, and maintenance.

Table 3. Recommender algorithms. Source: Microsoft, 2020b

Algorithm	Type	Supply Chain Application
Alternating Least Squares (ALS)	Collaborative Filtering	Matrix factorization for explicit or implicit feedback in integrated supply chain datasets
Cornac/Bayesian Personalized Ranking	Collaborative Filtering	Matrix factorization algorithm for predicting item ranking for technology and maintenance selection
Deep Knowledge-Aware Network	Content-Based Filtering	Deep learning incorporating a knowledge graph for technology/maintenance recommendations
Neural Collaborative Filtering (NCF)	Collaborative Filtering	Deep learning for implicit feedback from rich media data sets (forums, logs, chat)
Simple Algorithm for Recommendation	Collaborative Filtering	Similarity-based algorithm for evaluating similar technologies to the purchaser's selection
Surprise/Singular Value Decomposition	Collaborative Filtering	Matrix factorization for explicit rating feedback from small data sets (specialized purchases)
Vowpal Wabbit Family (VW)	Content-Based Filtering	Fast learning algorithm for scenarios where user features/context are constantly changing

THIS PAGE LEFT INTENTIONALLY BLANK

Prototype

We now advance a prototype for the acquisition supply chain management system, beginning with the initial requisition when the purchaser identifies the equipment required, the intended operations, the systems integration necessary (if any), and the required technical specifications (Figure 15).

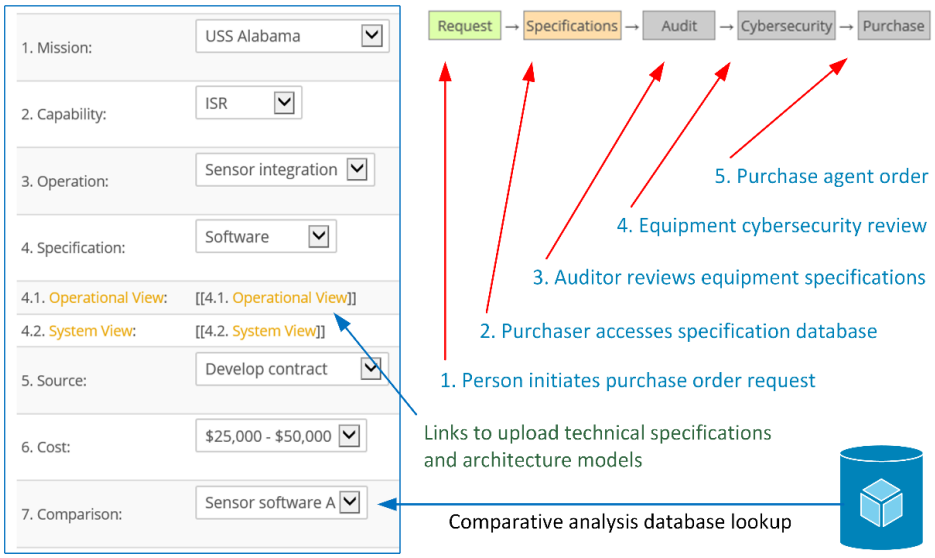

Figure 15. Section "A" pre-acquisition audit workflow

Here we build Section "A" from Figure 4, shown earlier in this report. The "source" determines whether this is to be a commercial purchase or builds upon an existing government capability. If software, whether to be developed in-house or outsourced. There is a "comparison" of similar systems—the first phase of the database lookup for similar products using the vector queries and recommender algorithms.

The in-service deployment audit (Figure 16) inspects the equipment immediately after purchase for counterfeit or substituted components and then inspects the technology in its operational configuration, including the impact of systems integration. The audit begins with initial deployment, progresses through installation and maintenance, to systems, networks, data, and cybersecurity audits, to the eventual

declaration of obsolescence and verification of destruction. Here we reference Section "B" from Figure 4.

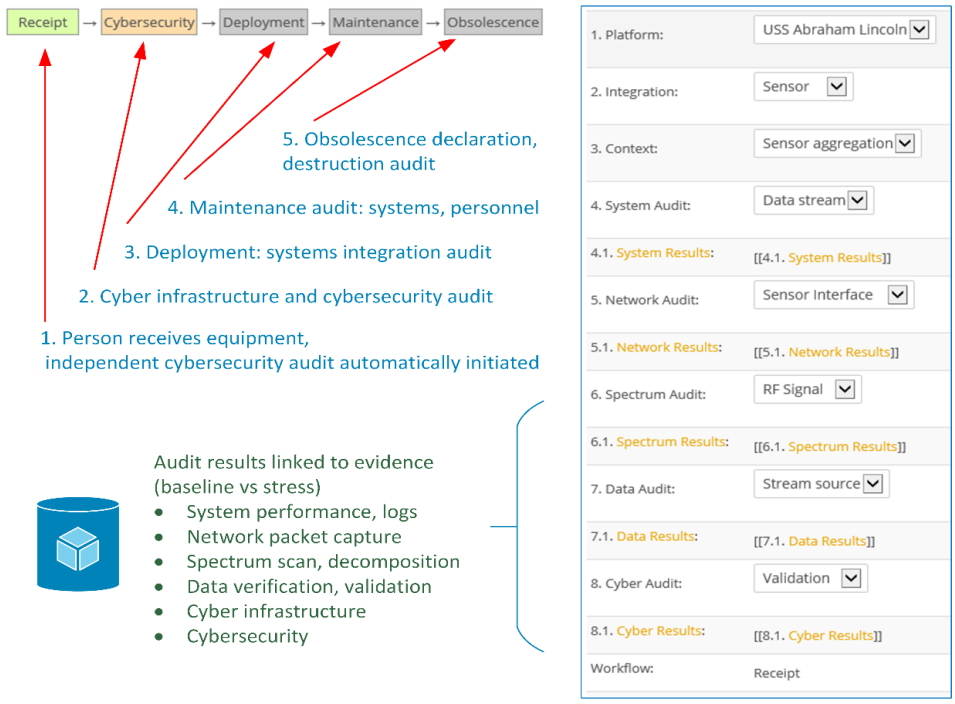

Figure 16. Section "B" in-service maintenance workflow

When the equipment is first received, the audit determines whether the supply chain has altered any of the components, software, or firmware. If the system does not meet the manufacturer specifications, or is without OEM components, then a refresh of that equipment is automatically initiated by the software.

The systems audit provides in-service technology performance measurements from independent auditors to address the performance of the host system within the technical, operational, and environmental contexts in which it is expected to be deployed. Audits address the systems under stress from cyber and electronic attack.

Network monitors assess inbound and outbound communications as well as network artifacts from systems integration. This includes the network interface on the hosting network, interconnections with supporting and supported equipment, adjacent equipment, and component dependencies.

Spectrum analysis assesses the RF environment on the system, measuring usage of the spectrum across system interfaces, bandwidth, and throughput of the system and its impact on the spectrum, and the impact of the environment on the systems from jamming, cyber, and electronic attack.

Data validation is critical. This involves mapping the data dependencies and data flows, and then measuring data transition and processing in the system in both baseline and compromised technical, operational, and environmental contexts.

Finally is the cyber audit in which we look collectively at the previous measures and then within systems and data for abnormalities or suspicious behaviors. The previously discussed AI algorithms are critical for analysis at this stage. Independent auditors should assess the overall cybersecurity posture of the systems environment and document the operation of the system and its data in different contexts. Then, of course the results of systems performance under simulated and actual cyberattack, and the results from failover and remediation.

Deployment status of the equipment across its operational life cycle is continuously assessed using one or more of the previously discussed regression algorithms. The auditors push the reports to the operations dashboards used by acquisition decision-makers.

THIS PAGE LEFT INTENTIONALLY BLANK

Conclusion

We presented a process for acquisition supply chain analytics, addressing data collection, tooling, and decision support. The workflow is largely automated such that reports for systems and components are presented to acquisition decision-makers across the life cycle of an asset. This understanding will help ensure cybersecurity compliance, systems viability, and data validity across the supply chain—from initial purchase request, to implementation and maintenance, to obsolescence.

Through the methodology, we build AI into the analytic process to help acquisition decision-makers manage the complexities of supply chain management including preemptive action to avoid eventual problems. Additionally, to automate the workflow and provide decision intelligence, we employ algorithms that integrate users, systems, and network information.

We presented predictive algorithms to help identify variables that may impact future acquisition decisions. The architecture integrated real-time event data within technical, operational, and environmental contexts. Future research may refine the algorithms or prototype online services to support acquisition supply chain analytics.

THIS PAGE LEFT INTENTIONALLY BLANK

References

Altman, N., & Krzywinski, M. (2015). Simple linear regression. *Nature Methods*, *12*(11), 999–1000.

Apache Spark. (2020, September 28). *Source code for pyspark.ml.recommendation*. https://spark.apache.org/docs/latest/api/python/_modules/pyspark/ml/recommendation.html#ALS

BMC. (2018a). *Beyond the hype—How do you really put AI to work for ITOps*. https://documents.bmc.com/products/documents/91/51/509151/509151.pdf

BMC. (2018b). *BMC Helix Client Management*. http://documents.bmc.com/products/documents/05/74/480574/480574.pdf

BMC. (2020, September 28). *Service-led operations: Re-imagine your service experience with the convergence of ITSM and ITOM*. https://www.bmc.com/it-solutions/service-operations.html

Carnegie Mellon University and The Johns Hopkins University Applied Physics Laboratory. (2020). *Cybersecurity Maturity Model Certification.*

Dalgaard, P. (2002). *Introductory statistics with R.* Springer-Verlag.

Department of Defense. (2019). *DoD digital modernization strategy.* Department of Defense.

Dietterich, T. (2003). Machine learning. In A. Ralston, E. Reilly, D. Hemmendinger, & (Eds.), *Encyclopedia of computer science* (pp. 1056–1059). Wiley.

Farsal, W., Anter, S., & Ramdani, M. (2018). Deep learning: An overview. *Proceedings of the 12th International Conference on Intelligent Systems* (pp. 1–6). Association for Computing Machinery.

Goldberg, D., Nichols, D., Oki, B., & Terry, D. (1992). Using collaborative filtering to weave an information tapestry. *Communications of the ACM*, *35*(12), 61–70.

Ixia. (2017). *Validating machine learning and security analytics*. Retrieved from Keysight: https://www.ixiacom.com/sites/default/files/2017-08/Ixia-S-PB-Machine-Learning-Analytics.pdf

Koren, Y., Bell, R., & Volinsky, C. (2008). Matrix factorization techniques for recommender systems. *Computer*, *42*(8), 30–37.

Maule, R. (2014). Cyber analysis grids for asynchronous distributed cloud services. *Proceedings of the International Conference on Grid & Cloud Computing and*

Applications. American Council on Science and Education. http://worldcomp-proceedings.com/proc/p2014/GCA2338.pdf

Maule, R. (2015). *Lifecycle methodology for system of systems engineering capability and integration analysis.* OPNAV.

Maule, R. (2016). Complex quality of service lifecycle assessment methodology. *Proceedings of the International Congress on Big Data* (pp. 462–469). IEEE.

Maule, R. (2017). *SEA cyber figure of merit tactical systems cybersecurity assessment* (NPS-N16-N116-A). Naval Postgraduate School. https://calhoun.nps.edu/handle/10945/57725

Maule, R. (2019a). Acquisition cybersecurity management framework. *Proceedings of the 16th Annual Acquisition Research Symposium* (pp. 1–16). Acquisition Research Program, Naval Postgraduate School.

Maule, R. (2019b). QoS for supply chain cyber management. *Proceedings of the World Congress in Computer Science, Computer Engineering, & Applied Computing* (pp. 3–9). American Council on Science and Education.

Maule, R. (2020). Cybersecurity audit framework for information assurance. *INFORMS Conference on Security* (pp. 1–21). Institute for Operations Research and the Management Sciences.

Maule, R., & Lewis, W. (2010). Security for distributed SOA at the tactical edge. *Proceedings of the Military Communications Conference* (pp. 13–18). IEEE.

Maule, R., & Lewis, W. (2011). Performance and QoS in service-based systems. *Proceedings of the World Congress on Services Computing* (pp. 556–563). IEEE.

Maule, R., Gallup, S., & Jensen, J. (2010). Knowledge engineering experimentation management system. In T. Sobh (Ed.), *Innovations and advances in computer sciences and engineering* (pp. 573–578). Springer.

McMullen, T. (2015). It probably works. *Communications of the ACM*, *58*(11), 50–54.

Microsoft. (2020a, September 22). *Deep learning vs. machine learning*. https://docs.microsoft.com/en-us/azure/machine-learning/concept-deep-learning-vs-machine-learning

Microsoft. (2020b, September 28). *Algorithm & module reference for Azure Machine Learning designer*. https://docs.microsoft.com/en-us/azure/machine-learning/algorithm-module-reference/module-reference

Microsoft. (2020c, September 28). *Recommenders*. Retrieved from Github: https://github.com/Microsoft/Recommenders

Microsoft. (2020d, September 28). *Smart Adaptive Recommendations (SAR) algorithm*. Retrieved from Github: https://github.com/Microsoft/Product-Recommendations/blob/master/doc/sar.md

Microsoft. (2020e, September 28). *Vector Search*. https://www.microsoft.com/en-us/ai/ai-lab-vector-search

Military & Aerospace Electronics. (2020). U.S. military authorities face replacing compromised chips in military computers. *Military & Aerospace Electronics, 31*(2), 9.

National Institute of Standards and Technology. (2020a, March 20). *Risk management framework*. https://www.nist.gov/cyberframework/risk-management-framework

National Institute of Standards and Technology. (2020b, September 23). *Cyber security framework*. https://www.nist.gov/cyberframework

NETSCOUT. (2019). *nGeniusONE Service Assurance platform*. https://www.netscout.com/sites/default/files/2018-12/EPDS_025_EN-1801-nGeniusONE.pdf

Oracle. (2020, September 28). Obtaining list of themes, gists, and theme summaries. *Oracle Text Application Developer's Guide*. https://docs.oracle.com/cd/A91202_01/901_doc/text.901/a90122/view3.htm

R. (2020, September 28). *The R project for statistical computing*. http://www.R-project.org/

Resnick, P., & Varian, H. (1997). Recommender systems. *Communications of the ACM, 40*(3), 56–58.

Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., & Riedl, J. (1994). GroupLens: An open architecture for collaborative filtering of netnews. *Proceedings of the Computer Supported Collaborative Work Conference* (pp. 175–186). Association for Computing Machinery.

Rich, C., Prasad, P., & Ganguli, S. (2019). *Market guide for AIOps platforms.* Gartner. https://www.gartner.com/doc/reprints?id=1-1XRR9HDN&ct=191115&st=sb

SPAWAR Office of the Chief Engineer. (2018). *Compile-to-combat in 24 hours (C2C24) implementation standard.* United States Navy Space and Naval Warfare Systems Command.

Unify. (2020, September 28). *Insights and discovery accelerator*. https://github.com/Azure-Samples/azure-search-knowledge-mining/blob/master/industry-solutions/documents/KnowledgeMiningWhitePaper.pdf

Villasenor, J., & Tehranipoor, M. (2013). The hidden dangers of chop-shop electronics: Clever counterfeiters sell old components as new, threatening both military and commercial systems. *IEEE Spectrum*. https://spectrum.ieee.org/semiconductors/processors/the-hidden-dangers-of-chopshop-electronics

Acquisition Research Program
Graduate School of Defense Management
Naval Postgraduate School
555 Dyer Road, Ingersoll Hall
Monterey, CA 93943

www.acquisitionresearch.net