04 Jan 2017

Asset Criticality in Mission Reconfigurable Cyber Systems and its Contribution to Key Cyber Terrain

Parker, Thomas

HICSS

Price, Peyton, et al. "Asset criticality in mission reconfigurable cyber systems and its contribution to key cyber terrain." Proceedings of the 50th Hawaii International Conference on System Sciences. 2017.

https://hdl.handle.net/10945/66971

Asset Criticality in Mission Reconfigurable Cyber Systems and its Contribution to Key Cyber Terrain

Peyton Price*, Nicholas Anthony Leyba*, Mark Gondree†, Zachary Staples*, Thomas Parker‡
*Naval Postgraduate School
nicholas.a.leyba.mil@mail.mil, zhstaple@nps.edu
†Sonoma State University
‡thomas.c.parker@navy.mil

***Abstract*—The concept of a common operational picture has been utilized by the military for situational awareness in warfare domains for many years. With the emergence of cyberspace as a domain, there is a necessity to develop doctrine and tools to enable situational awareness for key-decision makers. Our study analyzes key elements that define cyber situational awareness to develop a methodology to identify assets within key cyber terrain, thus enabling situational awareness at the tactical level. For the purposes of this work, we treat critical assets to be key cyber terrain, given that no formal study has determined differences between asset criticality and key cyber terrain. Mission- and operationally-based questions are investigated to identify critical assets with the TOPSIS methodology. Results show that the ICS system can be evaluated using TOPSIS to identify critical assets contributing to key cyber terrain, enabling further research into other interconnected systems.**

***Keywords*-key cyber terrain, cyber situational awareness, MADM, TOPSIS, industrial control system, cyber physical systems.**

I. Introduction

In the cyber warfare domain, the development of a common operational picture is necessary for the war fighter to appropriately deploy defensive countermeasures and offensive capabilities. Without situational awareness, the war fighter may make suboptimal decisions that will put people and equipment at risk. Current capabilities for cyber situational awareness are limited to network defense operations, such as intrusion detection, attack trend analysis, information flow analysis, damage assessment, and intrusion response [1]. To accurately attain cyber situational awareness, one must identify critical assets with respect to the mission and their importance to that mission. The criticality of individual assets and the relationship between assets describe the key cyber train for a given mission. Without identification of key cyber terrain, the war fighter suffers from reduced situational awareness and a diminished ability to defend and operate in the cyber domain.

Part of the difficulty in identifying critical assets in the cyber domain is in the dual logical and physical abstractions of the domain, the deep technical requirements for assessing asset behavior in a contested environment and the domain's complexity due to the growth of individual technologies into systems-of-systems [1]. Without asset or event filtering, interpreting the volume of data itself poses a deep technical challenge [2].

In support of this goal, our study validates an existing methodology for analyzing cyber asset criticality to determine key cyber terrain in the context of a non-trivial case study: a reconfigurable, shipboard industrial control system in use by the US Navy. We employ a hierarchical variant of TOPSIS, a multi-attribute decision making (MADM) strategy selected by Kim and Kang [3] as promising for defining cyber asset criticality. Our main contributions are as follows:

- We employ hierarchical TOPSIS in a complex setting, using the analytical hierarchy process for deriving weights for use in TOPSIS, exploring its *distinguishability* and *sensitivity* with respect to SME input;
- We "fill in the gap" for operationalizing this methodology, including defining cybersecurity criteria, decomposing a complex system into families and identifying missions;
- We explore criticality and key cyber terrain for a cyber-physical system for insight at the *tactical* level, whereas prior work focuses primarily on IT assets and the operational/strategic levels;
- We demonstrate that, as suspected, asset criticality is *mission-dependent* and contextual;
- In the context of our study, we find that weighting assets in relative importance to a mission plays a more important role in identifying cyber key terrain than ranking the relative importance of cybersecurity criteria.

Organization: In Sec. II, we discuss background and related work on identification of key cyber terrain. In Sec. III, we outline the methodology employed for our case study, a hierarchical variant of TOPSIS. In Sec. IV, we discuss the missions, criteria and assets of our case study. In Sec. V, we outline the factors explored in our case study and the results of our analysis. In Sec. VI, we conclude and discuss future work.

URI: http://hdl.handle.net/10125/41893
ISBN: 978-0-9981331-0-2

II. Background

The US Navy has prioritized the identification of key cyber terrain and cyber situational awareness in *U.S. Fleet Cyber Command / Tenth Fleet: Strategic Plan 2015–2020*. In particular, its Strategic Initiative 1.1 focuses the effort by stating:

> For each network ... and for each mission, we will 1) define key terrain, 2) identify or define operational availability (A_o) for that terrain, and 3) track how well we maintain A_o. This increased understanding will ensure we can successfully defend and fight through those key— and sometimes *decisive*— terrains. [4]

In our work, we treat asset crticiality as the sole measureable factor contributing to key cyber terrain. In the absence of work more concretely characterizing the relationship between these, we believe this is appropriate. We discuss existing efforts to describe key cyber terrain, analysis methods of determining critical cyber assets and related work.

A. Key Cyber Terrain

Traditionally, key terrain is defined as "any locality, or area, the seizure or retention of which affords a marked advantage to either combatant" [5]. In the cyber domain, key terrain involves network links and nodes that are essential to both friendly and adversarial forces [6]. What constitutes key cyber terrain has been disputed [7]–[11], but Franz [11] defines it as "the physical and logical elements of the [cyber] domain that enable mission essential war fighting functions; is temporal; ... and is applicable across strategic, operational, and tactical levels of war".

Raymond et al. [10] propose a framework for characterizing cyber terrain along the following planes: supervisory, cyber persona, logical, physical, and geographic. This framework partially aligns with DoD Joint Publication 3-12(R) [6] which depicts cyberspace into the three layers: the physical network layer, the logical network layer, and the cyber persona layer. While there are differences in identification of cyber domain planes, the two extra planes proposed by Raymond provide the ability to highlight command and control of cyber operations (supervisory plane) and tie in kinetic operations (geographical plane).

Jakobson [12] argues that key cyber terrain is made up of cyber assets and services, and their intra- and inter-dependencies. Cyber terrain is identified by focusing more on dependencies within a system vice across planes and is defined as three separate sub-terrains: hardware, which consists of a collection of connected network infrastructure components, like routers, servers, switches, communication lines, etc., and dependencies between them; software, which consists of different software components, such as operating systems and applications; and service, which represents all the services, such as database, file transfer, email, security services, etc., and their intra-dependencies [12].

Regardless of the framework or model utilized to identify key cyber terrain, a commonality in determining key cyber terrain is determining critical assets. MITRE emphasizes the importance of critical assets to key cyber terrain stating that "assets in operational environments are typically identified and their criticality determined via a mission impact analysis or business impact analysis" [8]. Dressler et al. [13] broaden the identification of key cyber terrain to include "all critical information, systems, and infrastructure; whether owned by the organization or used in transit by its information". From Franz' definition of key cyber terrain, MITRE and Dressler's discussions of critical assets as part of key cyber terrain tie directly into "mission essential war fighting functions" [11].

B. Asset Criticality

There is debate about the best method to determine asset criticality. Three proposed methodologies are attack graphs, Bayesian networks and multi-attribute decision making (MADM). Attack graphs are mathematical depictions of possible vectors of attacks against a specific network [14]. They can detail "all possible sequences of vulnerabilities an attacker can follow," or "with a monotonicity assumption stating an attacker never relinquishes an obtained capability, an attack graph can record the dependency relationships among vulnerabilities" [15]. Attack graphs are limited in that they quickly become complicated and require computer modeling, even for small networks [16]. Bayesian networks are graphical representations of probabilistic relationships within a network domain that can be constructed from attack graphs. They are limited by only accounting for a single actor and the assumed choices that the particular actor will make. If the assumed choices are incorrect, then a new Bayesian network is needed [17]. While these methods have value in looking at criticality from the attacker perspective or from the dependency aspect of asset criticality, MADM provides mature and well-understood multi-discipline methods for selecting the best decision among all feasible alternatives [3].

C. Related Work

Endsley originally proposed three key aspects for situational awareness [18], recently re-interpreted for the cyber domain by MITRE [19] as a framework comprised of network awareness (asset and configuration management), threat awareness (identifying incidents and suspicious behavior) and mission awareness (critical dependencies, real-time response, risk assessments and informed defense planning). MITRE's analysis focuses on the first two of these aspects, leaving the third area to human analysis. Our study is in the mission-centric analysis of cyber asset criticality and asset interdependencies, aiding the human analyst to prioritize assets for situational awareness. Jajodia et al. [20] reinforce both MITRE's and Endsley's work, observing that to "protect critical network infrastructures and missions,

we must understand not only the vulnerabilities of each individual system, but also their interdependencies and how they support missions" [20].

Several projects within the US Department of Defense seek to map cyber networks in a mission-centric manner—interpreting alarms and logs, monitoring processes, curating known vulnerabilities and SME input—with the goal of improving situational awareness. Schultz et al. survey these, calling them cyber network mission dependencies, or *mission mapping* [21]. Our apporach is complimentary to all of these, as every individual analysis methodology is insufficient in some aspect and most agree, ultimately, a hybrid approach is required for a comprehensive tool. In the terminology of Schultz et al., our methodology is a process-driven analysis as opposed to an artifact-driven analysis. Our approach, however, differs from existing process-driven analyses in terms of methodology and, thus, may provide useful cross-validation of their criticality determination or, otherwise, enhancing its coverage.

MIT Lincoln Lab's AMMO project uses SMEs to identify an initial set of assets comprising key cyber terrain and leverages network scanning increase terrain coverage. MITRE's CMIA project uses business modeling tools and uses SMEs to map assets to that model, to assess the impact of attacks to mission workflow during the system's design phase. MITRE's RiskMAP project employs a dependency-based analysis using SME input to model missions as a tasks and subtasks, for assessing the impact of attacks in disrupting mission goals in terms of availability, integrity and (unlike most network mapping tools) confidentiality. Johns Hopkins Applied Physics Lab's Dagger project is another dependency-based modeling tool using manual SME input to interpret impacts of cyber effects to missions. Unlike these tools, we employ a MADM methodology rather than a dependency-based analysis, the latter being a type of graph-based analysis. Criteria and asset weighting in MADM also permits a more flexible mechanism for incorporating SME judgement and asset inter-relationships, compared to explicit flow-based or dependency-based analysis.

III. Methodology

Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) is a MADM method that selects the best feasible alternative as the one closest to an *ideal* solution and farthest from the *negative ideal* solution. TOPSIS requires alternatives to have attributes which are monotonically increasing or decreasing, where all best attribute values comprise the *ideal* (zenith) solution and all worst values comprise the *negative-ideal* (nadir). We define a hierarchal variant of TOPSIS for a re-configurable system, S, relative to mission, M, given the following:

- attributes (criteria) $x_1, \ldots, x_n$
- alternatives (assets) $A_1, \ldots, A_m$
- score of A_i for attribute x_j is r_{ij}
- weight w_j for attribute x_j.

The goal is to find alternative A_i closest to *zenith* A^* and farthest from *nadir* A^-. Alternatives are grouped within a set of families for a particular mission (see Fig. 1). The scoring communicates the importance of that alternative within the mission as it pertains to each family. The *criticality score* for a given alternative is based on the highest score for the alternative over all the families.

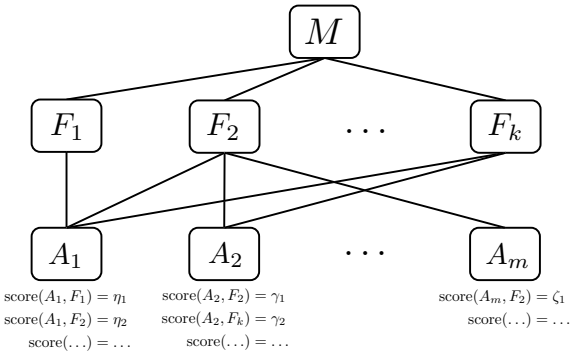

Figure 1. Scoring system assets based on attribute weights and allocation to families, relative to mission M.

A. Criteria Weighting

Feedback to weight attributes is solicited from experts, normalized and combined to get an overall weight following the Analytic Hierarchy Process (AHP). AHP allows "all important tangible and intangible, quantitatively measurable, and qualitative factors" to be included and measured [22]. We (i) solicit data to weight families (i.e., subsystems) with respect to each mission, (ii) solicit data to weight criteria irrespective of mission, (iii) solicit data to score each asset irrespective of mission and (iv) derive weights for criteria with respect to assets in a mission. The data is solicited from subject matter experts (SMEs) who have expert familiarity with the system and experience operating its subsystems. The system S is broken up into two or more hierarchical layers. For complex systems, division of S into parts reduces the need to gather feedback on weighting criteria per asset in each mission.

The system S is broken into subsystem component families $F_1, \ldots, F_k$. For each subsystem F_ℓ there is an associated set of assets, $\{A_{\ell,1}, \ldots, A_{\ell,m}\} \subseteq \{A_1, \ldots, A_m\}$. SMEs provide pair-wise comparisons among families expressing the relative importance of each to mission M. These comparisons are normalized via the eigenvector method (see Sec. III-B). This yields a set of normalized subsystem weights, $w_{a_1}, \ldots, w_{a_k}$, where w_{a_ℓ} is the weight for subsystem F_ℓ.

Independently, attributes, $x_1, \ldots, x_n$, are pair-wise compared by criteria subject matter experts with respect to their importance to system S irrespective of mission M. The

outcome of this process is a set of normalized attribute weights, $w_{b_1}, \ldots, w_{b_n}$, where w_{b_j} is the weight for criteria x_j.

Finally, weights for scoring the attributes $A_{\ell,1}, \ldots, A_{\ell,m}$ under subsystem F_ℓ are derived by multiplying subsystem weight w_{a_ℓ} by criteria weights w_{b_j}, yielding weights $w_{\ell,1}, \ldots, w_{\ell,n}$ for assets in subsystem F_ℓ during mission M. Concretely, for F_ℓ with $\ell \in [1, k]$, this is given by:

$$w_{a_\ell}\,(w_{b_1}, \ldots, w_{b_n}) = (w_{\ell,1}, \ldots, w_{\ell,n})$$

This relationship between weights within AHP is depicted in Fig. 2.

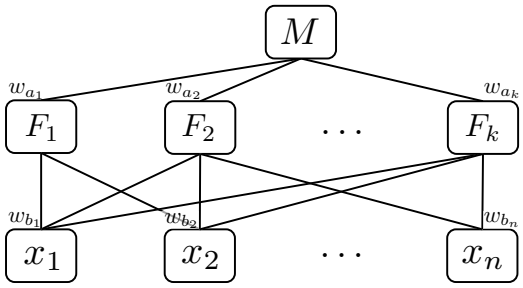

Figure 2. The role of weightings in the Analytical Hierarchy Process with subsystems and criteria.

B. Weighting Normalization

The eigenvector method allows for inconsistencies within pair-wise comparisons to be accounted for [22], [23]. The accomodated of matrix X is a non-null vector w such that $Xw = \lambda w$ and leaves w fixed. Accordingly, w is an eigenvector if it is a non-zero solution of $(X - \lambda I)w = 0$ for some λ. For an ideal solution with known weights, λ would be equal to the number of components n. This effect means that small variations within X are accounted for by keeping λ_{max}, the largest eigenvalue, close to n and the remaining eigenvalues close to zero [22, Ch. II, §5.1].

Next, we describe normalization in reference to attribute weights $w_{b_1}, \ldots, w_{b_n}$, but the process is the same for subsystem weights $w_{a_1}, \ldots, w_{a_k}$. First, the input weights $w_{b_1}, \ldots, w_{b_n}$ are normalized to have the property

$$\sum_{j=1}^{n} w_{b_j} = 1. \tag{1}$$

Next, we construct the reciprocal matrix X using the respondent's pair-wise comparison ratings, elicited via the AHP:

$$X = \begin{bmatrix} x_1/x_1 & x_1/x_2 & \cdots & x_1/x_n \\ x_2/x_1 & x_2/x_2 & \cdots & x_2/x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_n/x_1 & x_n/x_2 & \cdots & x_n/x_n \end{bmatrix}. \tag{2}$$

When multiplied by $\overline{w} = (w_{b_1}, \ldots, w_{b_n})$, the reciprocal matrix obeys the equation

$$(X - nI)\overline{w} = 0$$

since

$$X\ \overline{w} = \begin{bmatrix} 1 & x_1/x_2 & \cdots & x_1/x_n \\ x_2/x_1 & 1 & \cdots & x_2/x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_n/x_1 & x_n/x_2 & \cdots & 1 \end{bmatrix} \begin{bmatrix} w_{b_1} \\ w_{b_2} \\ \vdots \\ w_{b_n} \end{bmatrix} = nI\ \overline{w}.$$

Note, however, that X is not an exact measurement. Thus, we find the max eigenvalue, $\lambda_{\max}$, for the characteristic equation using:

$$\det(X - \lambda I) = \begin{vmatrix} 1-\lambda & x_1/x_2 & \ldots & x_1/x_n \\ x_2/x_1 & 1-\lambda & \ldots & x_2/x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_n/x_1 & x_n/x_2 & \ldots & 1-\lambda \end{vmatrix} = 0.$$

Using $\lambda_{\max}$, and given that the sum of the weights is one (see Eq. 1), the final weights $\overline{w}$ can found by solving the following homogenous system [23, Ch. 3]:

$$\begin{bmatrix} 1-\lambda_{\max} & x_1/x_2 & \ldots & x_1/x_n \\ x_2/x_1 & 1-\lambda_{\max} & \ldots & x_2/x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_n/x_1 & x_n/x_2 & \ldots & 1-\lambda_{\max} \end{bmatrix} \begin{bmatrix} w_{b_1} \\ w_{b_2} \\ \vdots \\ w_{b_n} \end{bmatrix} = 0.$$

C. Alternatives Selection

TOPSIS selects the most preferred alternative based on its closeness to the most preferred outcome (zenith) and distance from the least preferred outcome (nadir), using a Euclidean distance metric in the n-dimensional scoring space for criteria [3]. The proximity-to-zenith relationship can be modeled by:

$$C_{i^*} = \frac{A_{i^-}}{(A_{i^*} + A_{i^-})}. \tag{3}$$

First, the normalized decision matrix R must is created to transform each attributes score into a dimensionless metric for comparison across attributes [23, Ch. III, §2.3.5]. When y_{ij} is the scoring of asset A_i under criteria x_j, the decision matrix R is defined as:

$$R = \begin{bmatrix} r_{11} & r_{12} & \ldots & r_{1n} \\ r_{21} & r_{22} & \ldots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \ldots & r_{mn} \end{bmatrix}$$

where

$$r_{ij} = \frac{y_{ij}}{\sqrt{\sum_{i=1}^{m} y_{ij}^2}}.$$

Using decision matrix R and weights $\overline{w}$, we derive the weighted normalized decision matrix V, as follows:

$$V = \begin{bmatrix} v_{11} & v_{12} & \dots & v_{1n} \\ v_{21} & v_{22} & \dots & v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{m1} & v_{m2} & \dots & v_{mn} \end{bmatrix} = \begin{bmatrix} r_{11}w_1 & r_{12}w_2 & \dots & r_{1n}w_n \\ r_{21}w_1 & r_{22}w_2 & \dots & r_{2n}w_n \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1}w_1 & r_{m2}w_2 & \dots & r_{mn}w_n \end{bmatrix} = R\,\overline{w}\,.$$

The zenith, A^*, is the most preferable alternative available within the system and is comprised of the most desirable value of each asset over all the attributes. The nadir, A^-, is the least desirable and has the worst value of each asset over all attributes. The Euclidean distance from zenith and nadir is calculated for asset $i \in \{1, \dots, m\}$ as:

$$A_{i^*} = \sqrt{\sum_{j=1}^{n}(v_{ij} - v_j^*)^2} \quad \text{and} \quad A_{i^-} = \sqrt{\sum_{j=1}^{n}(v_{ij} - v_j^-)^2}$$

From A_{i^*} and A_{i^-}, the best feasible alternative C_{i^*} can be calculated from Eq. 3, and all the assets ranked by their relative closeness to the zenith [3].

IV. Case Study

Three key decision points for our study are defining the missions relative to which families are pair-wise compared, defining the criteria relative to which assets are scored, and defining the assets comprising the system S that are scored. We draw each mission M from existing doctrinal definitions used by the Navy to prepare, provide, and employ forces [24]. Criteria are derived from the Risk Management Framework (RMF) objectives outlined in Federal and Department of Defense security standards. Assets comprising system S are routers, and their associated connected subsystems, for a representative US Navy ship-board industrial control system.

A. Missions

The *Maritime Operations Center Standardization Manual* provides fleet commanders with an organization and process to bridge the gap between strategic guidance and tactical execution [24]. It outlines six essential mission areas and related supporting tasks. Our study selects three of these mission areas, relative to which scoring is performed and asset criticality is analyzed. These three missions, as defined by Joint Publication 1-02, are:

> Deterrence: The prevention of action by the existence of a credible threat of unacceptable counteraction and/or belief that the cost of action outweighs the perceived benefits.
>
> Sea Control: Employment of forces to destroy enemy naval forces, suppress enemy sea commerce, protect vital sea lanes, and establish local military superiority in vital sea areas.
>
> Power Projection: Conducting projection in the maritime environment to include a broad spectrum of military operations to destroy enemy forces or logistic support or to prevent enemy forces from approaching within enemy weapons range of friendly forces. [5]

Each mission may require different configurations of subsystems within S in order to successfully execute and accomplish that mission. These missions were chosen to provide three distinctly different system configurations.

B. Criteria

The criteria employed by our case study are given in Table I. Our criteria are based on the definitions of confidentiality, availability and integrity described in FIPS 199 [25] and draws impact terminology from NIST's Risk Management Framework [26]. Criticality needs to reflect the impact an attack may have on a compromised system with respect to those subsystems directly connected to the target and to subsystems that are indirectly connected (i.e., connected to the same asset within a family versus connected to another router within the family). Our study includes specific criteria for resource redundancy, reflecting the military's need to survive through asset nonavailability to continue operations.

C. Assets

The target system S is an ethernet-connected multiplex system comprised of multiple switches in a ring + 1 topology with multiple independent backbones (see Fig. 3). Routers are placed at the edge of the network with the backbones being made of switches. We limit our case study to the system's routers because they are the layer 3 devices connecting all major subsystem components in larger network.

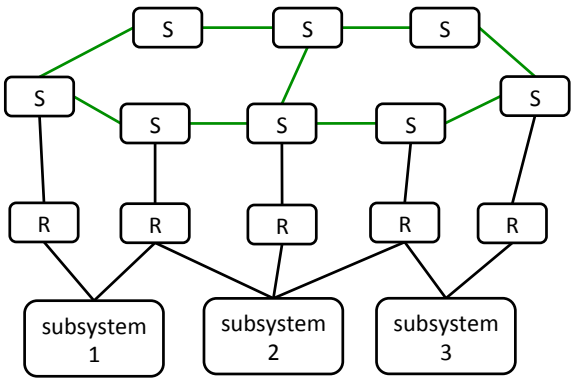

Figure 3. A notional topology resembling our target system, highlighting switches (S), routers (R) and subsystems.

Table I
SECURITY CRITERIA.

Criteria	Definition
Confidentiality to directly connected systems	If the target systems connected to the router were compromised, then the resulting unauthorized disclosure of information held by the target systems connected to that router could be expected to have *limited / serious / severe or catastrophic* adverse effect on organizational operations, organizational assets or individuals.
Integrity to directly connected systems	If the target systems connected to the router were compromised, then the resulting modification or destruction of information held by the target systems connected to that router could be expected to have *limited / serious / severe or catastrophic* adverse effect on organizational operations, organizational assets or individuals.
Availability to directly connected systems	If the target systems connected to the router were compromised, then the resulting disruption of access to or use of information held by the target systems connected to that router could be expected to have *limited / serious / severe or catastrophic* adverse effect on organizational operations, organizational assets or individuals.
Confidentiality to indirectly connected systems	If the target systems connected to the router were compromised, then the resulting unauthorized disclosure of information held by systems connected to other routers could be expected to have *limited / serious / severe or catastrophic* adverse effect on organizational operations, organizational assets or individuals.
Integrity to indirectly connected systems	If the target systems connected to the router were compromised, then the resulting unauthorized modification or destruction of information held by systems connected to other routers could be expected to have *limited / serious / severe or catastrophic* adverse effect on organizational operations, organizational assets or individuals.
Availability to indirectly connected systems	If the target systems connected to the router were compromised, then the resulting disruption of access to or use of information held by systems connected to other routers could be expected to have *limited / serious / severe or catastrophic* adverse effect on organizational operations, organizational assets or individuals.
Resource redundancy	There are *no / limited / multiple amount of* other systems to continue the same operation if the system connected to the router is compromised.

Each router has multiple subsystems connected to it. A subsystem deemed critical for a specific mission may carry information with a low priority but a high criticality. The categorization of systems in this manner allows for the efficient use of limited data, toward the task of identifying critical systems and assets within the ICS system with respect to mission and operational parameters.

Given a network map for S, each router for a subsystem can be scored against the criteria in Table I. The criteria are scored on a three point Likert scale, with one being *limited* and three being *severe or catastrophic*.

V. ANALYSIS AND RESULTS

In this section, we explore the results of applying our variant of TOPSIS methodology to our case study. As discussed in Sec. III, we solicit SME feedback to weight criteria, weight families and score assets. Typically, SME feedback collected via the AHP yields a single set of weights, derived via consensus. For our hierarchical variant, we solicit data from three different groups of SMEs (four criteria SMEs, four family SMEs, two asset SMEs) to derive a set of SME scores. By exploring this "scoring space," we intend to assess the relative impact of this data on the final results of the TOPSIS analysis and characterize the sensitivity of this methodology in determining key cyber terrain. In particular, it may be the case that one set of scores plays a much more important role than the others, and effectively decides key terrain in our case study.

For both family and criteria weighting, we expore three strategies: *random* weights, a randomly-generated, artificial weighting strategy; *average* weights, an average of all SME-derived weightings; and *transitive* weights, the single SME-derived weight demonstrating the best consistency ratio. As defined by Saaty [22], the consistency ratio is an expression of how strongly transitive relationships hold among weightings (i.e., $a < b$ and $b < c$ imples $a < c$), where highly transitive behavior yields ratios less than 0.1. We acknowledge, however, that inconsistency is not always bad for pair-wise comparisons as it reflects the complexity of real-world systems. For asset scoring, we explore two strategies: *informed* scoring, derived through SME consensus; and *uniform* scoring, an artificial score in which all assets are scored the same (i.e., two on the 1–3 Likert scale).

In total, fifty-four possible analyses (i.e., 3 missions, 3 family weighting strategies, 3 criteria weighting strategies and 2 asset scoring strategies) were compared to observe the sensitivity of these factors and how they impact the identification of key cyber terrain. For our analyses, we identify "good" results to be those that appear to be *highly distinguishable*, meaning they help establish distinguishing criteria highlighting assets of high relative criticality. All figures show more critical assets as having a higher "relative closeness" (per Eq. 3). Assets with high criticality are considered key cyber terrain. We leave exploring the possible difference between critical assets and key cyber terrain to future work.

Our observations are divided into three categories: the effects of family weighting, the effects of asset scoring and the effects of criteria weighting. Our preliminary analysis focuses on three factors: *impact*, or the effect of weightings and scoring; *transitivity*, or the effect of highly transitive weights versus average weights; and *perturbation*, or the effect of average weights versus random weights or, in

the case of asset scoring, informed scoring versus uniform scoring.

A. Impact of Family Weighting

For family weightings, we analyze the sensitivity of the results under small of changes in weights. We look at overall *impact* of family weights by observing the changes from differences in family weightings. For 48 of 54 analyses, the family weightings yield highly distinguishing results. The 6 analyses that we find not highly distinguishing are all within one mission (power projection) using non-random family weights; we suspect the relative closeness of family weights under that mission depressed differences in criticality. When analyzing *perturbation*, we compare pairs of analyses employing average or random family weightings, leading to 18 pairs of analyses. We observe, on average, 15 assets change in criticality[1], or approximately 50% of the total assets. This aligns with our expected behavior if assets changed their criticality following a coin toss, and shows that family weights indeed have an important role in the outcome of the analysis. In particular, if asset scoring dominates the outcome, we may not see this perturbation when asset scoring is held constant. In looking at *transitivity*, comparing highly transitive weightings (each time, using SME weightings with the lowest consistency ratio) to average weightings, leads to 18 pairs of analyses. Again, we compare changes in criticality. For example, comparing Figs. 4 and 5, we see three assets change in critical status and three change in non-critical status. On average, across each mission, we see between 4 and 9 assets change in criticality. Given that completely random changes in criticality yield 15 assets changing position, we consider a change in 4–9 assets to be quite sensitive to changes in family weightings.

In summary, we observe nearly all family weighting strategies were distinguishing, i.e., even random weights distinguish. This indicates that family weights *may* be inconsequential and that other weighting strategies dominate outcome; however, transitivity and perturbation analyses suggest that family weights are consequential and should be selected carefully. We observe that, given any fixed strategy for asset scoring and criteria weighting, changes in family weighting will perturb the key cyber terrain identified: for each of our missions, between 17.2–31.0% assets change position when comparing transitive vs. average strategies. This sensitivity suggests that the family weightings for a particular mission are highly distinguishable and highly influential for the methodology in discriminating asset criticality and key cyber terrain.

[1]In each analysis, assets are deemed either critical, non-critical or neither (exceeding a non-criticality threshold, but not exceeding a criticality threshold). We count assets that transition in their outcome: moving across these thresholds. Thus, an asset that moves from non-critical to critical would be counted as changing its criticality twice.

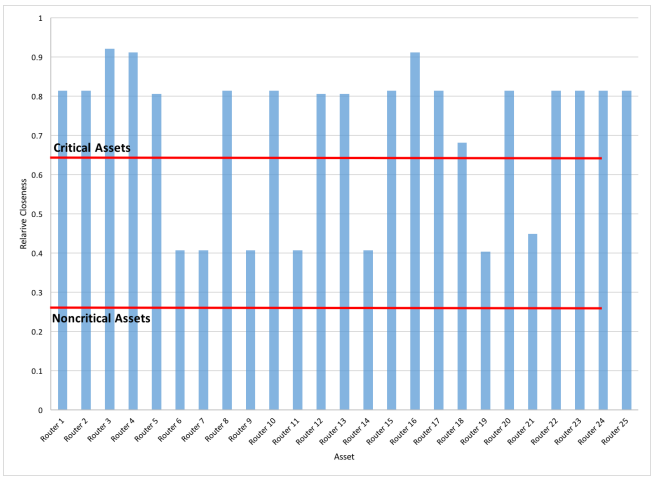

Figure 4. Asset criticality under the Deterrence Mission: average family weighting, transitive criteria, and informed asset scoring.

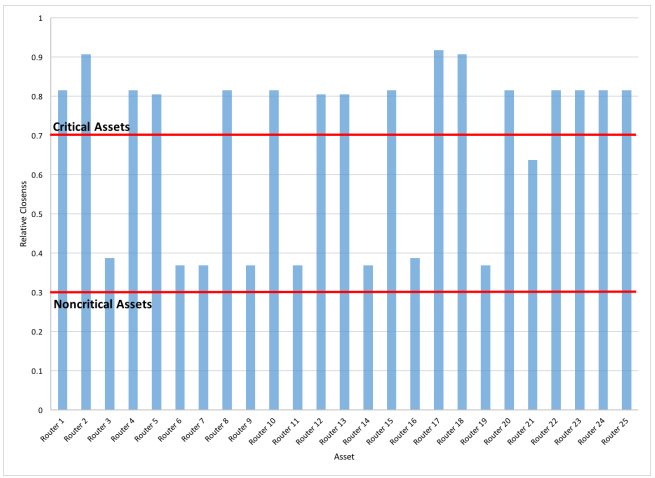

Figure 5. Asset criticality under the Deterrence Mission: transitive family weighting, transitive criteria, and informed asset scoring.

B. Impact of Asset Scoring

For assets, we analyze the *impact* and *perturbation* effects of scoring. When considering the impact of scoring, informed scoring results are highly distinguishable for 23 of 27 analyses. For perturbation analysis, we compared informed scores versus uniform scores: in 27 of 27 uniform scoring analyses, the effect of criteria weighting was largely minimized giving results that were predominantly based on the family weightings. For example, Fig. 6 and Fig. 7 demonstrate different per-mission family weights, which identify significantly different assets as critical. Varying scoring strategy and holding other factors constant leads to 27 pairs of analyses, comparing uniform to informed asset scoring. For 14 of these 27 pairs, uniform scoring results in more assets to be deemed critical; uniform scoring never resulted in fewer assets deemed critical, and never resulted in

more assets deemed non-critical. For example, considering a particular case within the power projection mission, six more critical assets were identified (compare Fig. 7 with Fig. 8). We conclude that uniform scoring produces more *conservative* results, which makes sense since it employs little judgement differentiating asset criticality. The value of having conservative estimates, potentially yielding a larger set of assets deemed critical, requires further study. Our study does not attempt to conclude if a more conservative approach to determining key cyber terrain is better or worse, from a practical application standpoint.

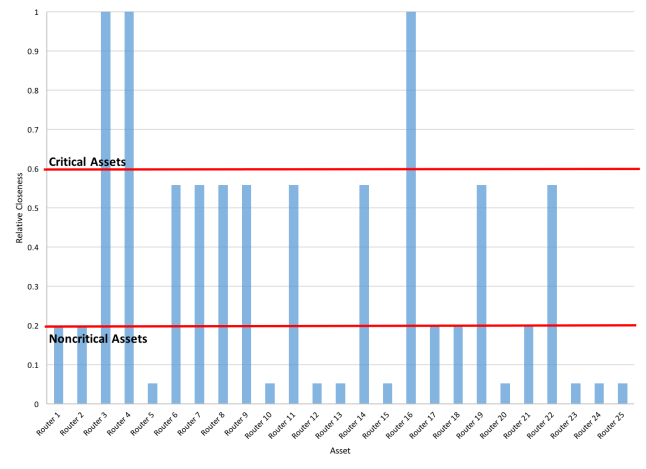

Figure 6. Asset criticality under the Sea Control Mission: average family weighting, transitive criteria, and uniform asset scoring.

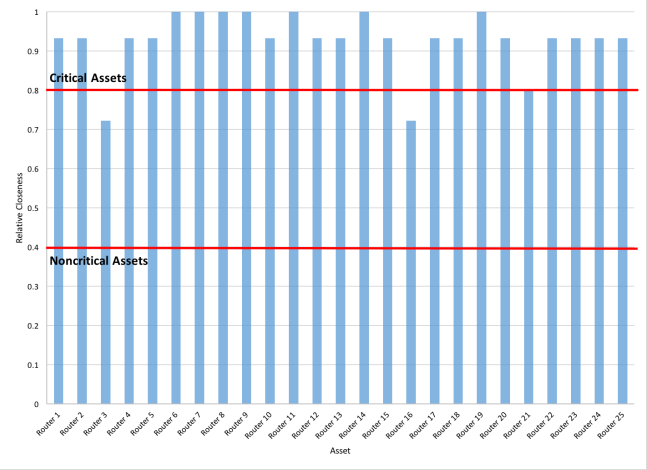

Figure 7. Asset criticality under the Power Projection Mission: average family weighting, transitive criteria, and uniform asset scoring.

C. Impact of Criteria Weighting

The sensitivity of results to criteria weights was analyzed to determine effect on key cyber terrain. We explore the

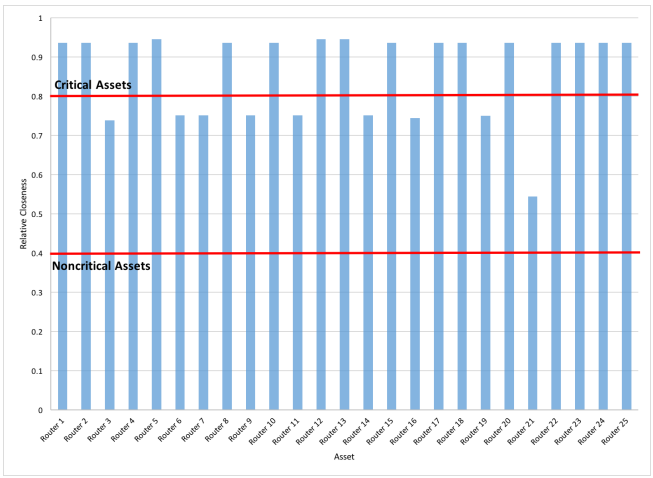

Figure 8. Asset criticality under the Power Projection Mission: average family weighting, transitive criteria, and informed asset scoring.

effect of *perturbation* for criteria weighting, comparing random with average criteria weighting—while holding other factors constant, using only non-random family weighting and informed asset scores[2]—yields 6 pairs of analyses (i.e., three missions, two family weightings, informed scoring). Across these 6, we observe an average of 2 assets change criticality (6.9% of all assets); excluding one case (power projection mission with average family weightings under informed scoring), this average is significantly lowered: 0.83 assets (2.9% of all assets). Thus, either outcomes exhibit low-sensitivity to criteria weights, or average criteria weights are effectively random. In fact, we do observe high variance among SME-derived weighting: the largest difference between derived single criteria weightings is 0.3027. We expected, given our use of AHP, to be unable to achieve *total* consensus on what security criteria are most important for key cyber terrain; but there were few similarities in criteria weighting to suggest even the general importance of each criteria.

We find similar trends when analyzing the effect of *transitivity* for criteria weighting—i.e., employing the most transitive weights while holding other factors constant, as before—again, yielding 6 pairs of analyses. Across these 6 comparisons, we observe an average of 0.5 assets change criticality (1.7% of all assets), demonstrating that outcomes show low-sensitivity to transitivity in criteria weighting. This extremely low effect with highly transitive weights suggests that outcomes are more significantly decided by family weighting and asset scoring strategies, and see little difference among highly transitive, average or random criteria weighting strategies. Qualitatively, however, it appears

[2]We exclude random family weighting and uniform asset scoring, as we have previously remarked the effects of these, i.e., random changes to asset criticality or more conservative determination of asset criticality, respectively.

that highly transitive weights yield simpler to interpret distinguishing criteria among critical and non-critical assets; we leave fuller characterization of this observation to future work.

Given the apparent low sensitivity of criteria weights, we explore the magnitude of effect of criteria weights compared to family weights. We conduct our analyses by observing the change in the number of critical assets when the family weighting strategy is changed and criteria weighting strategy is held constant (i.e., average, transitive, and random criteria weighting strategy under informed scoring when family weighting is most transitive or average), and the number of critical assets when the criteria weighting strategy is changed and family weighting strategy is held constant (i.e., average, transitive, and random family weighting strategy under informed scoring when criteria weighting is most transitive or average). This leads to two sets of 18 analysis pairs: in one set, family weighting strategy is varied and in the other, criteria weighting strategy is varied. When looking at the family weighting strategies, we observe an average of about 7 assets change criticality (23.7% of all assets); when criteria weighting strategies change, we observe an average of less than one asset changes criticality (1.1% of all assets). As illustrated in Figs. 4 and 5, six assets changed criticality when family weighting strategy changed. In comparison, as shown in Figs. 4 and 9, only two assets changed criticality when criteria weighting changed[3]. This large difference in number of assets changing criticality suggests that family weightings have about 20 times larger effect on the outcomes in identifying key cyber terrain than do criteria weighting in the context of our case study.

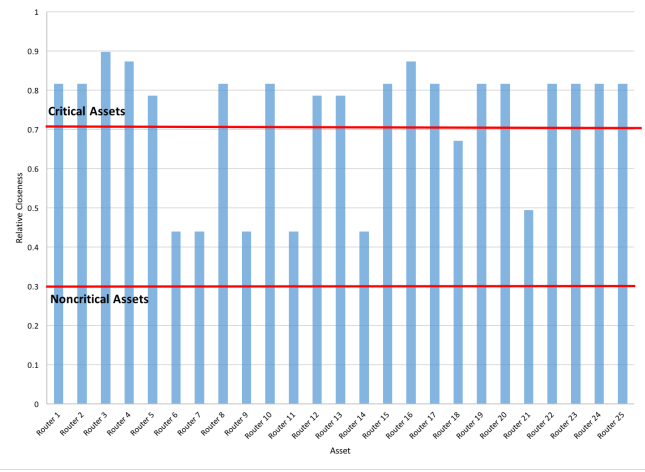

Figure 9. Asset criticality under the Deterrence Mission: average family weighting, average criteria, and informed asset scoring.

[3]Note: This particular comparison had the largest number of critical asset changes when the criteria weighting strategy changed. Only one other comparison had any assets change criticality with the remaining comparisons having zero changes in critical assets.

VI. Conclusion

Our application of hierarchical TOPSIS yields insight into the cyber-physical key terrain aboard a US Navy vessel required for the successful conduct of various missions. The methodology used here suggests that critical assets' participation in key cyber terrain can be identified and does change according to mission. The impact of our findings means that cyber situational awareness should account for mission. As a result, we conclude that any cyber situational awareness tool must account for the changes in the operational use of a system while operating under different missions.

Within the context of our case study, our observations suggest the most important factor in determining key cyber terrain is in defining family weights for a particular mission. Once per-mission family weights are determined, then the key cyber terrain derived from the methodology used here will closely mirror those weights. Our observations suggest that criteria weights are dramatically less important than family weights. When assets are scored uniformly, key cyber terrain appears to be identified in a more conservative and less informed manner.

An important corollary of our research in practical application is that identifying key cyber terrain will allow scarce resources to be prioritized to critical systems during operation. Further work is needed to demonstrate that terrain prioritization benefits situational awareness because it allows focus on the health and performance of key assets without distraction from system components that do not impact mission performance. Similarly, further work may show that terrain prioritization allows the use of tailored analytics and limited defensive resources to focus on monitoring and protecting infrastructure essential to the success of the ship.

A. Future Work

Further work is required to validate the soundness of the results obtained using this methodology, i.e., that assets identified as contributing to key cyber terrain are the same as those identified as integral in attack and vulnerability studies for the target system. This could be done by overlaying known vulnerabilities with critical assets to better identify the possible impact to missions and effect on key cyber terrain, or compare our study with other methodologies to determine the differences in identification of key cyber terrain. It would be interesting to explore the high SME-response variance we observed. In fact, no formal survey work has been conducted to elicit family and criteria weightings, so no reserch exists to guide selecting a sample size to achieve statistical significance. Exploring how to elicit pairwise rankings with confidence in the statistical properties of the derived weights would be a valuable contribution, allowing future research to avoid the costly process of achieving consensus-derived weights.

REFERENCES

[1] P. Barford, M. Dacier, T. G. Dietterich, M. Fredrikson, J. Giffin, S. Jajodia, S. Jha, J. Li, P. Liu, X. Ou, D. Song, L. Strater, V. Swarup, G. Tadda, C. Wang, and J. Yen, *Cyber Situational Awareness: Issues and Research*. Boston, MA: Springer, 2010, ch. Cyber SA: Situational awareness for cyber defense, pp. 3–13.

[2] W. W. Streilein, J. Truelove, C. R. Meiners, and G. Eakman, "Cyber situational awareness through operational streaming analysis," in *IEEE Military Communications Conference (MILCOM 2011)*, Nov. 2011, pp. 1152–1157.

[3] A. Kim and M. H. Kang, "Determining asset criticality for cyber defense," Office of Naval Research, Tech. Rep. NRL/MR/5540–11-9350, 2011.

[4] *U.S. Fleet Cyber Command/TENTH Fleet: Strategic Plan 2015–2020*, U.S. Fleet Cyber Command, Washington, DC, 2015. [Online]. Available: http://www.navy.mil/strategic/FCC-C10F\%20Strategic\%20Plan\%202015-2020.pdf

[5] *Joint Publication 1-02: Department of Defense Dictionary of Military and Associated Terms*, JP 1-02, Department of Defense, Washington, DC, 2010. [Online]. Available: http://www.dtic.mil/doctrine/new_pubs/jp1_02.pdf

[6] *Joint Publication 3-12 (R): Cyberspace Operations*, JP 3-12, Department of Defense, Washington, DC, 2013. [Online]. Available: http://www.dtic.mil/doctrine/new_pubs/jp3_12R.pdf

[7] P. W. Phister, "Cyberspace: The ultimate complex adaptive system," *The International C2 Journal*, vol. 4, no. 2, 2010. [Online]. Available: http://www.dodccrp.org/files/IC2J_v4n2_03_Phister.pdf

[8] D. Bodeau, R. Graubart, and W. Heinbockel, "Mapping the cyber terrain: Enabling cyber defensibility claims and hypotheses to be stated and evaluated with greater rigor and utility," MITRE, McLean, VA, Tech. Rep. MTR130433, 2013.

[9] J. R. Mills, "The key terrain of cyber," *Georgetown Journal of International Affairs*, pp. 99–107, 2012. [Online]. Available: http://www.jstor.org/stable/43134343

[10] G. Conti, T. Cross, M. Nowatkowski, and D. Raymond, "Key terrain in cyberspace: Seeking the high ground," in *2014 6th International Conference on Cyber Conflict*. Tallinn: NATO CCD COE, 2014. [Online]. Available: http://www.westpoint.edu/acc/SiteAssets/SitePages/Publications/06916409.pdf

[11] G. J. Franz III, "Effective synchronization and integration of effect through cyberspace for the joint warfighter," presented at Armed Forces Communications and Electronics Association TechNetLand Forces-East, Baltimore, MD, August 2012. [Online]. Available: http://www.afcea.org/events/tnlf/east12/documents/4V3EffSynchIntEffthruCybrspcforJtWarfighter_forpublicrelease.pdf

[12] G. Jakobson, "Mission-centricity in cyber security: Architecting cyber attack resilient missions," in *Proceedings of the 5th International Conference on Cyber Conflict (CyCon)*. IEEE, June 2013, pp. 1–18.

[13] J. Dressler, W. Moody, C. L. Bowen III, and J. Koepke, "Operational data classes for establishing situational awareness in cyberspace," in *Proceedings of the 6th International Conference On Cyber Conflict (CyCon 2014)*, June 2014, pp. 175–186.

[14] P. Ammann, D. Wijesekera, and S. Kaushik, "Scalable, graph-based network vulnerability analysis," in *Proceedings of the 9th ACM Conference on Computer and Communications Security*. ACM, 2002, pp. 217–224.

[15] L. Wang, S. Noel, and S. Jajodia, "Minimum-cost network hardening using attack graphs," *Computer Communications*, vol. 29, no. 18, pp. 3812–3824, 2006.

[16] R. E. Sawilla and X. Ou, "Identifying critical attack assets in dependency attack graphs," *ESORICS*, pp. 18–34, 2008.

[17] P. Xie, J. Li, X. Ou, P. Liu, and R. Levy, "Using Bayesian networks for cyber security analysis," in *IEEE/IFIP International Conference on Dependable Systems and Networks*, 2010, pp. 211–220.

[18] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems," *Human Factors and Ergonomics Society*, vol. 37, no. 1, pp. 32–64, 1995.

[19] MITRE. (2016, January) Situation awareness: Today's leaders need meaningful cyber situation awareness to safeguard sensitive data, sustain fundamental operations, and protect national infrastructure. [Online]. Available: http://www.mitre.org/capabilities/cybersecurity/situation-awareness

[20] S. Jajodia, S. Noel, P. Kalapa, M. Albanese, and J. Williams, "Cauldron: Mission-centric cyber situational awareness with defense in depth," in *IEEE Military Communications Conference (MILCOM 2011)*, 2011, pp. 1339–1344.

[21] A. E. Schulz, M. C. Kotson, and J. R. Zipkin, "Cyber network mission dependencies," MIT Lincoln Laboratory, Tech. Rep. 1189, May 2015.

[22] T. L. Saaty, *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*. New York: McGraw-Hill Book Co., 1980.

[23] C.-L. Hwang and K. Yoon, *Multiple Attribute Decision Making : Methods and Applications a State-of-the-Art Survey*, M. Beckmann and H. P. Kunzi, Eds. New York: Springer-Verlag, 1981.

[24] *Maritime Operations Center Standardization Manual*, OPNAV M-3500.42, Department of the Navy, Washington, DC, 2014.

[25] *Federal Information Processing Standards Publication 199: Standards for Security Categorization of Federal Information and Information Systems*, FIPS 199, National Institue of Standards and Technology, 2004.

[26] *National Institute of Standards and Technology Special Publication 800-37: Guide for Applying the Risk Management Framework to Federal Information Systems: A Security Life Cycle Approach*, NIST SP 800-37, National Institute of Standards and Technology, February 2010.