

Chapter 2

Overview of Cyber Situation Awareness

George P. Tadda and John S. Salerno

Abstract Improving a decision maker's¹ situational awareness of the cyber domain isn't greatly different than enabling situation awareness in more traditional domains². Situation awareness necessitates working with processes capable of identifying domain specific activities as well as processes capable of identifying activities that cross domains. These processes depend on the context of the environment, the domains, and the goals and interests of the decision maker but they can be defined to support any domain. This chapter will define situation awareness in its broadest sense, describe our situation awareness reference and process models, describe some of the applicable processes, and identify a set of metrics usable for measuring the performance of a capability supporting situation awareness. These techniques are independent of domain but this chapter will also describe how they apply to the cyber domain.

2.1 What is Situation Awareness (SA)?

One of the challenges in working in this area is that there are a multitude of definitions and interpretations concerning the answer to this simple question. A keyword search (executed on 8 April 2009) of 'situation awareness' on Google yields over 18,000,000 links the first page of which ranged from a Wikipedia page through the importance of "SA while driving" and ends with a link to a free internet radio show. Also on this first search page are several links to publications by Dr. Mica Endsley whose work in SA is arguably providing a standard for SA definitions and

George P. Tadda and John S. Salerno, Air Force Research Laboratory Rome NY

¹ Decision maker is used very loosely to describe anyone who uses information to make decisions within a complex dynamic environment. This is necessary because, as will be discussed, situation awareness is unique and dependant on the environment being considered, the context of the decision to be made, and the user of the information.

² Traditional domains could include land, air, or sea.

techniques particularly for dynamic environments. In [5], Dr. Endsley provides a general definition of SA in dynamic environments:

“Situation awareness is the perception of the elements of the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future.”

Also in [5], Endsley differentiates between situation awareness, “a state of knowledge”, and situation assessment, “process of achieving, acquiring, or maintaining SA.” This distinction becomes exceedingly important when trying to apply computer automation to SA. Since situation awareness is “a state of knowledge”, it resides primarily in the minds of humans (cognitive), while situation assessment as a process or set of processes lends itself to automated techniques. Endsley goes on to note that:

“SA, decision making, and performance are different stages with different factors influencing them and with wholly different approaches for dealing with each of them; thus it is important to treat these constructs separately.”

The “stages” that Endsley defines have a direct correlation with Boyd’s ubiquitous OODA loop with SA relating to Observe and Orient, decision making to Decide, and performance to Act. We’ll see these stages as well as Endsley’s three “levels” of SA (perception, comprehension, and projection) manifest themselves again throughout this discussion.

As first mentioned, there are several definitions for SA, from the Army Field Manual 1-02 (September 2004), Situational Awareness is:

“Knowledge and understanding of the current situation which promotes timely, relevant and accurate assessment of friendly, competitive and other operations within the battlespace in order to facilitate decision making. An informational perspective and skill that fosters an ability to determine quickly the context and relevance of events that are unfolding.”

From [2](pg120):

*“When the term **situational awareness** is used, it describes the awareness of a situation that exists in part or all of the battlespace at a particular point in time. In some instances, information on the trajectory of events that preceded the current situation may be of interest, as well as insight into how the situation is likely to unfold. The components of a situation include missions and constraints on missions (e.g., ROE), capabilities and intentions of relevant forces, and key attributes of the environment.”*

The components of a situation that Alberts identifies are key points of analysis that can be performed as a situation is being recognized or identified. Of particular interest are an analysis of capability, opportunity, and intent when considering friendly or competitive players. [2](pgs 18-19) also provides a definition for awareness as:

“Awareness exists in the cognitive domain. Awareness relates to a situation and, as such, is the result of a complex interaction between prior knowledge (and beliefs) and current perceptions of reality. Each individual has a unique awareness of any given military situation.”

And a separate definition for understanding:

“Understanding involves having a sufficient level of knowledge to be able to draw inferences about the possible consequences of the situation, as well as sufficient awareness of the situation to predict future patterns.”

Alberts clarifies these three definitions with (note the parallel with Endsley’s concepts of perception and projection):

“Hence, situation awareness focuses on what is known about past and present situations, while understanding of a military situation focuses on what the situation is becoming (or can become) and how different actions will impact the emerging situation.”

The distinction between Alberts and Endsley is that Alberts separates awareness and understanding while Endsley includes understanding (projection) as a part of awareness. Alberts also seems to imply that analysis of the situation can only be performed as cognitive processes. Finally in [13], the authors don’t specifically address SA but they do define situation and impact assessment (recall that Endsley drew a distinction between situation awareness and situation assessment but concluded that situation assessment enables situation awareness):

“Level 2 - Situation Assessment: estimation and prediction of relations among entities, to include force structure and cross force relations, communications and perceptual influences, physical context, etc.”

“Level 3 - Impact Assessment: estimation and prediction of effects on situations of planned or estimated/predicted actions by participants; to include interactions between action plans of multiple players (e.g. assessing susceptibilities and vulnerabilities to estimated/predicted threat actions given one’s own planned actions)”

There are many other definitions available for situation awareness but the ones described above seem to be or are becoming widely accepted. The JDL Data Fusion Model has been used since the 1990’s to describe ideas for sensor fusion and multi-sensor fusion, Dr. Endsley’s work over the past 21 years has been used to define several SA supporting applications, and Alberts’ work has become the defining work for Network-Centric Operations and Warfare.

Then, What is Situation Awareness? Let’s use Endsley’s definition with a slight modification as first published in [10]:

*“Situation awareness is the perception of the elements of the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future to **enable decision superiority.**”*

This ties together the definitions given above if we allow Alberts's definition to apply to computer automation and to provide additional analysis of the perceived reality. We also need to equate his definition of understanding with Endsley's definition of projection which is a natural association given how they both consider anticipating the future given the current situation. Note that all these definitions include the element of time. Time involves the use of past experience and knowledge to identify, analyze, and understand the current situation and the projection of possible futures. These enable a decision maker to maintain awareness, make decisions, and take action to influence the environment which then requires an update to the situation, causes more decision and actions, and results in a continuous cycle. Reference [3] describes this decision cycle but additional detail is outside the scope of this chapter. The time dimension and continuous nature are also what cause the environment to be dynamic since the situation and elements within the situation will change as time progresses.

The remainder of this chapter will present a model that provides a reference for the above definition of SA, a supporting process model, a breakdown of some of the components and concepts captured by the models, and finally some measures of performance and effectiveness. Most of the ideas described can be applied to any domain and can be considered domain agnostic. Section 4 contains an example for how the domain independent ideas can be specifically applied to the cyber domain.

2.2 Situation Awareness Reference and Process Models

This section extracts from [8], [9], [11], [14] and [15] specific details about the reference model and definitions used for the reference model. The section concludes with a development of a generic process model.

2.2.1 Situation Awareness Reference Model

According to Endsley, SA begins with perception. *Perception* provides information about the status, attributes, and dynamics of relevant elements within the environment. It also includes classifying information into understood representations and provides the basic building blocks for comprehension and projection. Without a basic perception of important environmental elements, the odds of forming an incorrect picture of the situation increase dramatically. *Comprehension* of the situation encompasses how people combine, interpret, store, and retain information. Thus, comprehension includes more than perceiving or attending to information; it includes the integration of multiple pieces of information and a determination of their relevance to an individual's underlying goals and can infer or derive conclusions about the goals. Comprehension yields an organized picture of the current situation by determining the significance of objects and events. Furthermore, as a

dynamic process, comprehension must combine new information with already existing knowledge to produce a composite picture of the situation as it evolves. Situation Awareness refers to the knowledge of the status and dynamics of the situational elements and the ability to make predictions based on that knowledge. These predictions represent a **Projection** of the elements of the environment (situation) into the near future.

McGuinness and Foy [6] extended Endsley's Model by adding a fourth level, which they called **Resolution**. This level tries to identify the best path to follow to achieve the desired state change to the current situation. Resolution results from drawing a single course of action from a subset of available actions. McGuinness and Foy believe that for any fusion system to be successful, it must be resilient and dynamic. It must also address the entire process from data acquisition to awareness, prediction and the ability to request elaboration (drill-down) for additional data and finishing with an appropriate action. McGuinness and Foy put Endsley's model and their model into perspective with an excellent analogy. They state that Perception is the attempt to answer the question "What are the current facts?" Comprehension asks, "What is actually going on?" Projection asks, "What is most likely to happen if...?" And Resolution asks, "What exactly shall I do?" The answer to the resolution question isn't to tell a decision maker what specific action to perform or what specific decision to make but instead provides options of end actions and how they affect the environment. Specifics about the actual decision or course-of-action to execute to achieve a chosen effect are carried out by command and control functions.

Another point to be made is that any proposed model should not promote a serial process, but rather a parallel one. Each function (for example in Endsley's model: Perception, Comprehension, Projection with the added Resolution) happens in parallel with continuous updates provided to and from each other. It should also be emphasized that each sub-component (in both models) also continuously interacts with each other and embarks its data/knowledge to the others. Another important note is that throughout any analysis each step should provide a high level of visibility or transparency to the decision maker.

Our SA Reference Model, shown in 2.1, is built by combining the JDL Data Fusion model and Endsley's SA Model. In addition to presenting the model, definitions of the various components of the model are provided. In particular, we've refined how one can think of JDL Levels 1 and 2 as well as describe differences between JDL Levels 2/3 and Endsley's idea of projection.

There continues to be a debate as to what JDL Levels 1 and 2 represent. One belief is that JDL Level 1 deals only with the tracking and identification of individual objects while JDL Level 2 is the aggregation of the objects into groups or units through the identification of relationships between the objects. For example, JDL Level 1 objects could be various equipments (tanks, APCs, missiles, etc). At JDL Level 2, equipment along with personnel can be aggregated into a unit or division based on time and space. JDL Level 1 attempts to answer such questions as Existence and Size Analysis (How many?), Identity Analysis (What/Who?), Kinematics Analysis (Where?), and includes a time element (When?). But if we consider this separation between JDL Levels 1 and 2 then several questions arise; how do we

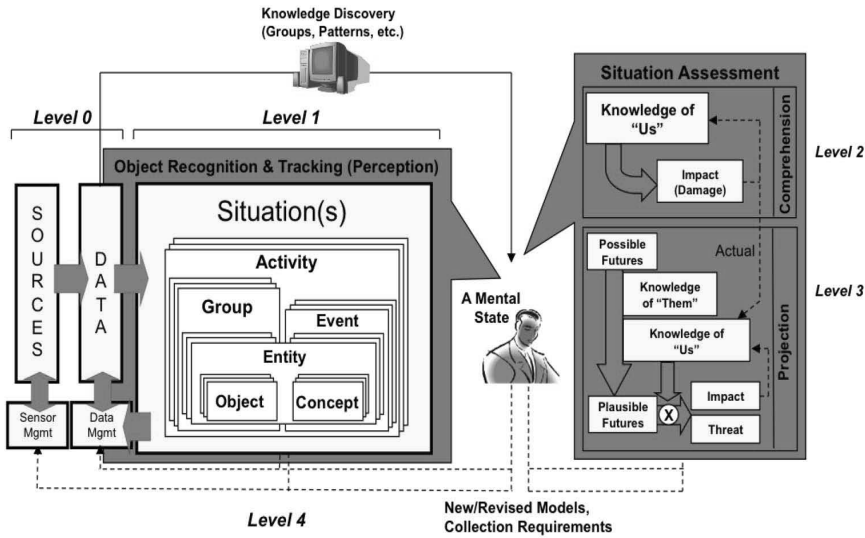


Fig. 2.1 Situation Awareness Reference Model

account for concepts or non-physical objects and can't we track a group or activity like an object? What is a situation? How does the system acquire the necessary a priori knowledge (or relationships) to perform aggregation? What is the difference between models for identifying an object, a group, or an activity?

To begin to answer these questions we first present a number of basic definitions and then use them to refine what we mean by JDL Level 1 and 2. We then will explore the difference between JDL Levels 2 and 3 and what Endsley refers to as Projection.

In [1]³, an **entity** is defined as "something that has a distinct, separate existence, though it need not be a material existence. In particular, abstractions and legal fictions are usually regarded as entities. In general, there is also no presumption that an entity is animate. The word entity is often useful when referring to something that could be a human being, a non-human animal, a non-thinking life-form such as a plant or fungus, a lifeless object, or even a belief." An **object** is "a physical entity; something that is within the grasp of the senses" [1]; "something perceptible by one or more of the senses, especially by vision or touch" (The Free Dictionary). What if the entity is not a physical object? How can we describe it? Generally speaking, an abstract entity still can be associated with a time or existence and an abstract concept (e.g., a phone call, financial transaction, etc.)

A **group** is "a number of things being in some relation to each other". A group can be an interest group (terrorist cell, religious order) or an organizational

³ Clearly we don't want to use Wikipedia® as a definitive reference. However, references to it in this chapter are simply for definitions, e.g., as a dictionary.

group (police, government, Non-Governmental Organization (NGO) or military). An **event** is “something that takes place; an occurrence at an arbitrary point in time; something that happens at a given place and time” [1]. Both entities and groups can be associated with a specific event or events. Snidaro, Belluz and Foresti [12], further decompose an event into 3 classes: Simple, Spatial, and Transitive. They define a Simple event as one which involves only a single entity and with no interaction with other entities; a Spatial event describes events that occur in space and include a location. The third event, Transitive, involves two entities that are connected by some interaction. Spatial events can be a tank (entity) or unit (group) being at a given location at a specified time. If the tank or unit then interacts with another tank or unit then we say we have a Transitive event.

An **activity** is “something done as an action or a movement” [1]. Activities are composed of entities/groups related by one or more events over time and/or space. Thus, by definition an event, group or activity can be considered a complex entity (or in terms of the JDL, an object) and can be tracked and identified similarly to a simple entity. As a side note, the JDL Lexicon, [7], defines an entity as “Any object or object set (or event or event set) which forms the basis of a hypothesis used in data fusion processes” but does not provide a definition for an object or event.

By using the definitions presented above, we argue that activities and the set of these activities at a point in time (which we refer to as the situation) is both a part and a result of JDL Level 1. Models or a priori knowledge is necessary for JDL Level 1 to be capable of identifying the object, group or activity. This a priori knowledge (i.e., the relationships or associations) can be learned through Knowledge Discovery and validated by an operator or provided directly. Here we note that Knowledge Discovery techniques only learn statistically relevant occurrences. As such, new or novel ideas cannot be learned and require knowledge elicitation or conjecture of possible existence by a human. Actually, the conjecture of an activity by an experienced decision maker is a key activity or set of activities of interest that must be considered when providing capabilities for SA. So what are examples of activities? Classical activities can range from force-on-force actions of a conventional war, potential multi-stage or coordinated cyber attacks, potential terrorist attacks (asymmetric) to operations other than war. These activities are composed of a number of interconnected and inter-related events and processes.

We define a **situation** as a *person’s world view of a collection of activities that one is aware of at an instance in time*. We also argue that a computer system can identify an activity is occurring based on some a priori knowledge and interconnect a number of objects/events but cannot itself develop or provide Situation Awareness; only a person (the decision maker) can be aware. A computer is a tool that can assist/support a person in developing and maintaining awareness. Thus we argue that there is one situation or world view per person based on their context. **Shared Situation Awareness** is then a consensus view of a number of individual views about a specific activity or set of activities. Likewise, there is a growing community supporting “Shared action plans” to represent group decision making over jointly observed information or data reduction.

The JDL Level 2 definition (given in the first section of this chapter) does not distinguish between time, current or future, while JDL Level 3, Impact/Threat Assessment is specifically associated with the future (estimate of ‘predicted’ actions). Why can’t we have a current threat or impact? How is the current situation different from the projected or forecasted one? Can we have different impacts/threats depending on the timeframe that we are projecting? Our research is leading us to look at JDL Level 2 or Endsley’s comprehension level as addressing the current situation (assessment of the current situation tends to be *damage assessment* since the impacts or effects have already occurred) and looking at JDL Level 3 and Endsley’s projection level as the projection of the current situation and its analysis (i.e., future impacts and threats). Thus, we split the assessments represented by JDL Levels 2 and 3 based on time rather than functionality.

Additionally, Bosse, Roy and Wark [4] define **Situation Assessment** as, “**a quantitative evaluation of the situation that has to do with the notions of judgment, appraisal, and relevance.**” Two products or components of situation assessment are: Impact and Threat Assessment. Impact assessment is defined as:

“...the force of impression of one thing on another; an impelling or compelling effect. There is the notion of influence: one thing influencing another. In that sense, impact assessment estimates the effects on situations of planned or estimated/predicted actions by the participants, including interactions between action plans of multiple players.”

In [4], they also define threat assessment as “*an expression of intention to inflict evil, injury, or damage. The focus of threat analysis is to assess the likelihood of truly hostile actions and, if they were to occur, projected possible outcomes...*” The only difference we note in their definition of impact/threat assessment is that, like the JDL definition, they are only concerned with the future.

Based on the definitions in [4] we can further define situation assessment as the understanding of the current situation and what it means to ‘me’ (its damage), the projection of the current situation into the future (which we refer to as the set of plausible futures) and the potential impacts/threats of those plausible futures. In Endsley’s Level 2, or comprehension, we need to have an understanding of “us” and what is important to “us” (commonly referred to in the literature as “Blue” but can also include “Grey”). In order to accomplish this we need to know such information as to our resources (capacity and capabilities), what is important to us (salience) and what our vulnerabilities are. Based on this information the identified activities within the current situation can be ranked based on their impact (associated damage) and threat (increased/decreased). Or, the activities could be ranked based on most likely (the greatest impact) and most dangerous (the greatest threat).

Feedback in any control system is very important, especially in an ever changing and dynamic environment. Here we discuss what type of feedback or what JDL calls Process Refinement (Level 4) means for JDL Levels 2/3 and conceptually how it can be implemented. We also present how it is affected by Projection. The basic definition of **Process Refinement** covers two separate but integrated capabilities. For the purpose of our discussion we will divide them into external and internal process.

Externally, we are concerned with providing sensors or collections with positioning information based on forecasted or anticipated movement of objects/entities or groups. The classical example here is the tracking of an object. A common tracking algorithm used in today's system is a Kalman Filter. Kalman Filters provide the ability to forecast where the object could be in one time increment in the future. This position information can then be provided to "better" position the sensor. Theoretically, a similar approach can be done with concepts and groups to include non-physical entities.

Recall that our revised definition of Level 2 is concerned with assessment of the current situation. As one develops their understanding of the current situation, questions may arise and more data could be required to either fill in holes or reduce the uncertainty of given data. These requirements can be considered as additional or revised *collection requirements* and be provided as feedback to the collection requirements process. Level 3 can provide similar data to the collection process, except from a somewhat different perspective. Projected activity or activities are just that and from a single current situation multiple futures can be developed. From each of these futures, the analyst can determine key events that could assist them in determining which one is unfolding. These key events can be used to drive the collection requirements process.

Internal processes also need to be monitored to ensure that the information processing system is performing as designed. At the object level one can suggest, possibly based on environmental inputs, which source is "better" at that time for tracking or identifying the object or sending the same sensor data to multiple algorithms (running in parallel), coming up with possibly different answers and combining the results in some manner. Similar concepts can be used at the activity level. Additionally, a priori knowledge or models could be updated as a part of an internal process refinement function. As new information comes in and new knowledge is developed through the analysis and projection processes, a decision maker (or analyst) may update existing models or add/create new models. It is important to understand that tools can be provided (e.g., data mining, knowledge discovery, etc.) that can assist the user in finding new relationships or patterns but in many cases they also produce meaningless patterns or noise. These tools tend to also be based on past behaviors and activities and may not produce meaningful models to identify new behaviors or activities. In such a case it is important that the human use these tools as input and verify/validate the results. Such tools typically cannot come up with "novel" or never before seen patterns or relationships for which there is little or no data (or in most cases not statistically relevant) to support it. The human is still by far the most capable to develop such models and any technique/interface must take this into account. Figure 2.1 provided a graphical conception of how each of the components defined in this section fit together.

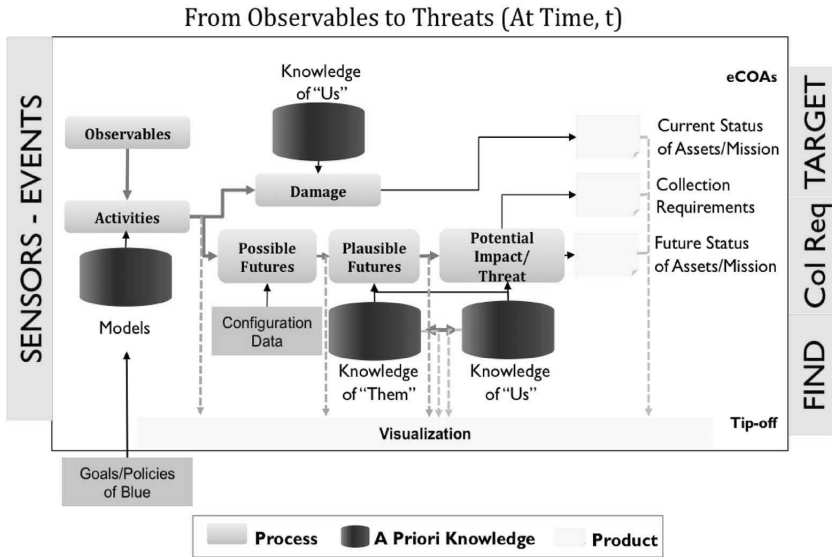


Fig. 2.2 Situation Awareness Process Model

2.2.2 Situation Awareness Process Model

2.2 expands upon the reference model and looks at it as a process in an instance of time. Observables are the input to the process that provides a view of what is going on in the world (primitive elements of the environment). It is assumed that any attributes associated with the observables have been normalized, cleansed, and transformed into a form that can be used by the follow-on processes⁴. The observables we are interested in are cues into the activities that a decision maker needs or is interested in (and thus we refer to these as Activities of Interest, AOI) as a way to gain or maintain awareness. The AOI are based on goals, policies, or in general the “things” of interest. These AOI can be stored and manipulated in such formats as graphs, Bayesian networks, Markov models, or any of the numerous modeling techniques. As observables enter the process, they are categorized and (1) associated with a new stage or step within an existing, ongoing activity; (2) associate with no existing activity and hence become the start of a new activity; or (3) can be a trigger leading to the combination, merging, or removal of existing activities. This process is similar to tracking individual objects (generally referred to as Level 1 fusion) and why we consider this part of the process, even though mostly symbolic, still a Level 1 process or as a form of object ID and tracking. However, in this case our object is an activity, a complex object. The classical tracking problem of association also

⁴ The specific techniques to perform the operations of normalization, cleansing, or transformation are out of scope of this discussion.

comes in to play when associating an observable to an activity or step of an activity. These activities could also be thought of as hypotheses.

At any given time, say t , we have a set of ongoing activities (defined earlier as the current situation). At this point we are interested in analyzing the meaning of these activities? This is considered to be Situation Assessment (as shown on the right side of the reference model in Figure 2.1). The overall objective of Situation Assessment is to determine if any of the ongoing activities have an impact to ‘us’⁵ or if they can have (in the future) an impact to ‘us’. The first part looks at the current activities and assessing the impact that the activities have had. Since these activities have already happened, we refer to this as “Damage” Assessment, i.e., has any of the identified activities caused an impact and specifically has it caused harm that requires development of a recovery plan to resolve? In order to accomplish this type of assessment, one not only needs the current, known activities, but also what it means to ‘us’. The information needed is part of what we call “Knowledge of Us”. This data contains knowledge as to the importance of the assets or capabilities to ‘us’. Thus, this part of the process identifies to the decision maker whether there is a current impact (damage) to any of our capabilities or assets used to perform a mission as well as any current impacts (damage) to the assets themselves.

Above, we discussed the current situation and assessing the situation including its impact to the mission, but a decision maker may also be interested in a view of what the adversary (or competitor) is doing or may possibly do. This has generally been described as “getting inside the adversary’s OODA loop”. The sooner we understand what the adversary can/might do the more options become available to the decision maker. The second part of 2.2 addresses this. The first step of the process is to take each of the activities of interest and project them forward based on the a priori knowledge provided as part of the model. Here we don’t discuss time itself, i.e., we are not projecting the activities based on time, but rather the next step. In some cases it could take milliseconds to go from one stage to another and in other cases it could be days, or longer. The number of stages that we look forward is defined under “Configuration Data”. So based solely on the models themselves, we have projected each current activity one step forward; however, these projected or possible futures do not take into account whether they are plausible. In order to determine plausibility, we need to consider additional knowledge. We need both the “Knowledge of Them” and “Knowledge of Us”. Specifically, we need to know whether the adversary has the capability, capacity, and intent/goal and have they possibly exhibited similar behavior in the past. We also need to know whether they have the opportunity to accomplish the intent(s)/goal(s). This opportunity is based in many cases on the vulnerabilities of us (provided as part of the Knowledge of Us). Thus, starting with the list of possible futures, we use the “Knowledge of Them” and “Knowledge of Us” to constrain the possible into the plausible for each activity of interest.

But again what do these plausible futures mean to me/us? To answer this question we again use part of the “Knowledge of Us” (importance of the assets/capabilities)

⁵ ‘Us’ is considered to be friendly assets or the friendly environment. In a military sense, ‘us’ is what we’re interested in defending. In a philosophical sense, ‘us’ can be equated to self and the preservation of self. Or, it can be thought of as the defensive environment of the decision maker.

to identify potential impacts and threats to meeting our objective(s). From this portion of the process we get not only future potential impacts/threats but we can also use this knowledge to determine our future collection requirements. Based on each of the futures, we can identify the key differentiating events that will assist us in determining which of the futures are actually unfolding. The key differentiating events can then determine the collection requirements needed to increase the certainty in identifying whether a plausible future is occurring.

One of the dangers in a reference model such as the one in 2.1 is that it can be perceived as a sequential flow of data or information rather than a descriptive model of components and ideas. To help circumvent this danger, 2.2 attempts to define a process flow and end products that is based on the concepts of the reference model as its framework. A primary feature of the process model is that it defines components that can be implemented as automated computer applications or shared human/computer systems that can then be tied together within system architectures. It also describes the flow of information and when key data sources come into play.

This section has described two models and defined their components. The SA Reference Model provides a set of definitions that can serve as a reference for describing systems that aid with SA while the Process Model captures a process flow at a single instance of time. Together, the two models provide a common set of definitions for situation awareness. One component of both models not yet addressed is Visualization. This will be the topic of the next section.

2.3 Visualization

Visualization is a component across all aspects of both models. Almost any part of the process is open to visualization but comes with its own set of challenges. Depending on what aids the decision maker and in what context they need awareness, activities of interest can be visualized, as well as results of various assessments to include current and future impacts. However, quickly and completely conveying the situation (set of AOIs) to a decision maker, especially in light of large amounts of information, is a very challenging issue.

There are three primary challenges to visualization for this research area. The challenges come about when dealing with abstract or conceptual elements as opposed to strictly geo-spatial or physical elements and result in challenges in visualizing the situation, high-volume data domains, and rapidly conveying the situation and corresponding analysis to a decision maker.

First is the challenge of visualizing a situation without necessarily depending on geo-spatial displays or displays that depend on physical objects. There are a multitude of techniques that are used to visualize raw or roughly correlated data but conveying situations has proven to be a challenge. What is the best way to present both the current and future situation so that sufficient decisions can be reached? There is also the added complexity of visualizing the situation over time so that changes, trends, and projections to the situation can be conveyed. In some contexts,

a simple geo-spatial display may be sufficient but when addressing abstract concepts as opposed to physical entities, a geo-spatial display is only a part of the information that needs to be presented to a decision maker. Situation visualization remains a significant research challenge.

In any type of visualization, the ability to reach back into the raw data and to be able to follow the processing path is very important. For a user to be comfortable with the analysis a system performs, a certain level of trust in the conclusions need to be established. This trust can develop through transparency throughout any automated processes and through allowing the user to explore the data themselves to decide if they'd reach the same conclusions.

2.4 Application to the Cyber Domain

The ideas described in the previous sections are not dependent on any particular domain and can be directly applied to the cyber domain as well as many other domains or areas. There are however many challenges in actually implementing them as a way to enable situation awareness for a decision maker. Key to this approach is the identification and modeling of the relevant activities of interest. These activities are the principle driver for what a fusion engine would look for in the observables presented to it. The activities become the basis for the analysis and assessments done to determine current impacts of assets and missions and to derive plausible futures and their future impacts or effects on assets and missions.

What activities of interest might be useful in the cyber domain? To date, this work has focused primarily on defining a model of complex multi-stage network attacks. As a matter of fact, work on the detection and evaluation of this activity has driven a large portion of the research in this area but the work has so far been limited to the single activity of attacks. In exploring other potential AOIs, a network administrator might be interested in policy change detection, firewall configuration problems, network attack, etc. Cyber activities of interest for a company manager might include misuse of company equipment, effect of attacks on business units, etc. There also may be a need for awareness of cross domain activities as we move from operational to more a more strategic level of decision maker. An Air Operations Center commander may be interested in how an effect in the cyber domain influences the ability to execute a flying mission; a dispatcher may be interested in how a network attack influences the routing of city taxis, etc. The point of this is that by using a model of an activity of interest there is tremendous flexibility in how the activity is defined for the area that a decision maker needs awareness and they aren't limited to a single domain at a time. But the challenge is in defining that activity of interest and the observables that can be used to detect or identify that activity.

Specifically for the cyber domain, we've defined a single model to address the AOI of a complex multi-stage cyber attack. The observables are based on common Intrusion Detection Systems as well as application and firewall logs. The current situation is based on the set of "attack tracks" identified at the current time. We're

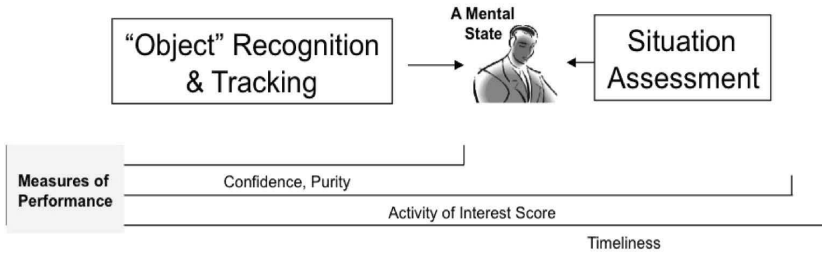


Fig. 2.3 Metrics Mapped to SA Reference Model

currently working on technologies to identify the impact of each attack track on the network assets affected by the attack. In the case of cyber, the environment is the network topology and the collection of data associated with that topology. This environmental information becomes the “knowledge of us” which can be used in performing impact assessment of the current situation. As our research progresses, we’re researching the construction and interaction of each of the components within Figure 2.

2.5 Measures of Performance and Effectiveness

This section extracts from [8], [14], and [15]. [8] describes metrics in a generic sense while [14] and [15] are specific to the cyber domain. For the cyber domain, the metrics are specialized by substituting ‘attack track’ (an attack track is defined as a hypothesis of a complex multi-stage cyber attack that contains all the evidence of an attack) for ‘activity of interest’. The current set of generic measures of performance have four dimensions; confidence, purity, cost utility, and timeliness. The metrics that will be described using the definitions for the models described in Section 2. Those definitions describe the various sections of the model with input streams representing “sources”; the output from network sensors and other evidence representing “data”; and attack tracks corresponding to activities and situations. 2.3 shows, at a high level, how each of the metrics that will be discussed map onto the reference model.

The metrics currently measure a system’s ability to correctly fuse evidence, produce attack tracks, and prioritize the attack tracks into a meaningful order for a user. One of the metrics, *attack score*, has applicability to systems capable of assessing impact/threat and should improve in value (get closer to 1.0) as a system can more completely analyze the attack track in the context of a network or in importance to effects on a mission. Attack score is in the process of being updated to measure more information than just attacks. The updated measure is being called an “activity of interest score”. One last comment on metrics before getting into the bulk of the discussion concerns data reduction. Early work on measures, spoke at length about

a Data-Information Ratio (DIR) as shown in equation 2.1. The DIR was intended to measure the overall reduction in the amount of “stuff” that was presented to a user. When data is presented, a user tends to be: 1) overwhelmed by volume and lack of context; 2) has to rely on individual expertise for understanding; and, 3) has to mentally process (fuse, assess, and infer) the data.

$$DIR = \frac{\text{Number of Complex Entities}}{\text{Number of Observations}} \quad (2.1)$$

The initial thought was that by automatically aggregating and organizing data into more useful information the user would have less to deal with and could more productively and more effectively maintain awareness of the environment. The DIR has proven to be very informative but at a fairly high level. It tends to indicate the capability of a class of work rather than the capability of individual systems. For instance, in the cyber domain, we’ve observed an on average data reduction of two orders of magnitude when processing “alerts” (observations or events) into attack tracks (complex entities or activities). The advantage to the user then is that instead of tens of thousands of individual pieces of data to consider they now only have to consider a few hundred possible attack tracks. An attack track only reduces the information initially presented while maintaining the ability to “drill down” into the more detailed data that makes up the track. When combined with a mechanism to prioritize importance of the possible attacks, the power of this general class of analysis begins to become apparent. However, beyond the general data reduction, the DIR doesn’t provide a lot of insight into the performance of particular technologies or implementations.

The rest of this section will discuss each of the metrics in more detail. The metrics, or measures of performance, are discussed according to the four dimensions; confidence, purity, cost utility, and timeliness. As shown by the mapping in 2.3, the confidence and purity measures address object recognition and tracking, cost utility includes situation and impact assessment and timeliness covers the entire model.

2.5.1 Confidence

For the cyber domain and assuming our activity of interest is a cyber attack on a network, **confidence** is a measure of how well the system detects the true attack tracks (e.g., the hypothesis of an attack). The confidence dimension consists of four metrics; (1) recall, (2) precision, (3) fragmentation, and (4) mis-association. Consider the diagram shown in 2.4 it represents the space of attack tracks identifiable by a capability supporting Cyber SA. The attack tracks can be classified into three categories; (1) known tracks, (2) detected tracks, and (3) correctly detected tracks. *Known tracks* are the attack tracks given by ground truth and contain all the evidence for a particular attack. *Detected Tracks* are the attack tracks hypothesized by a Cyber SA system under evaluation and contain all the fused evidence. Finally, *Correctly Detected Tracks* are known attack tracks detected by the Cyber SA system. Known

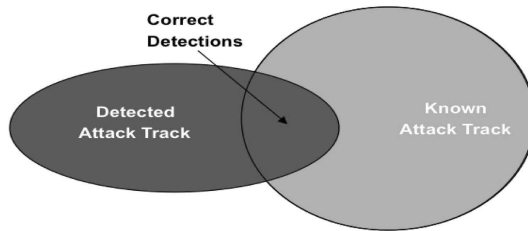


Fig. 2.4 Space of Attack Tracks

tracks not detected would be false negatives and detected tracks not known would be false positives.

Given this description, the confidence metrics are as described by the equations below:

$$Recall = \frac{Correct\ Detections}{Known\ Attack\ Tracks} \quad (2.2)$$

$$Precision = \frac{Correct\ Detections}{Detected\ Attack\ Tracks} \quad (2.3)$$

$$Fragmentation = \frac{Number\ of\ Fragments}{Detected\ Attack\ Tracks} \quad (2.4)$$

Misassociation =

$$\frac{Number\ of\ Detections\ that\ are\ neither\ correct\ detections\ nor\ fragments}{Detected\ Attack\ Tracks} \quad (2.5)$$

Two additional definitions are needed to complete the discussion of confidence metrics. A *fragment* is said to be an attack track that should have been included within another track (equation 2.4). For example, in an island-hopping attack, a targeted computer is compromised and then used to originate attacks on other computers. To correctly detect this attack, any time a target becomes an attacker, all the evidence of this attack should be included in same track as that from the original attacker through the island all the way to the subsequent targets. Often a fusion engine will not correctly associate the subsequent evidence with the original attack reporting them as two or more attack tracks. In other words, a single attack is reported as two or more attacks only one of which would be counted as the correctly detected track. A fragment can give the appearance of a false positive when, in reality, it usually indicates a portion of a more complex attack. *Mis-association* is a little simpler and captures all “other” detected tracks. A mis-associated track is one which is neither a fragment nor correctly detected. Summed together, the values for precision, fragmentation, and mis-association should sum to 1 or cover 100% of the

attack tracks produced by the Cyber SA system. Assessing Cyber SA systems under research and development showed the confidence metrics to be the most useful when determining the overall capability of a system thus far. The traditional tension between recall and precision was observed in that high recall usually meant low precision-detected lots of attacks but couldn't clearly identify them. While, in contrast, high precision (knew the exact attack type and details) would often result in low recall or missed attacks. The fragmentation metric was interesting from the perspective of complex attacks and the level of data reduction. High fragmentation led to more attack tracks being presented which lowered data reduction but still typically maintained the two orders of magnitude reduction. The greatest value in the fragmentation metric was in indentifying that it would be better to keep more complex attacks in a single attack track and also in generating discussion concerning attribution of an attack. For example, if there was evidence of two attack tracks from different original attackers attacking the same target, *A*. Then, there was subsequent evidence that *A* was attacking *B*. Which of the original two attackers would be attributed to the follow-on attack on *B*? How would the attribution be made? Could you be sure that one of those attackers continued on and that it wasn't a new attack originating with *A*? These questions remain open areas of research.

2.5.2 Purity

Purity characterizes the quality of the correctly detected tracks (again assuming our activity of interest is a cyber attack on a network). Purity metrics also “look into” a track at the evidence and provide indications as to how well the evidence is being correlated and aggregated into an attack track. There are two metrics used to measure purity; (1) Mis-assignment rate, and (2) Evidence Recall. The equations for these metrics are given below:

$$\text{Mis-AssignmentRate} = \frac{\text{Total No. Alerts in Results} - \text{Total No. Correct Alerts for Assigned Results}}{\text{Total No. Alerts in Results}} \quad (2.6)$$

$$\text{EvidenceRecall} = \frac{\text{Total No. Correct Alerts for Assigned Results}}{\text{Total No. Alerts in Ground Truth Identified}} \quad (2.7)$$

By looking at the quality of the correctly detected tracks, the thought was that the metric could indicate how well the Cyber SA system was using the available evidence. Mis-assignment rate could answer the question about whether the system was assigning evidence to a track that wasn't relevant or if it only considered directly useful evidence. While evidence recall was intended to tell us how much of the evidence available was truly being used. When applying the purity metrics

to the cyber domain, neither proved to be particularly useful. Mis-assignment rate was probably the stronger of the two in that when the rate was very high it would indicate an incorrect correlation or association of the underlying data. This would essentially indicate a “bug” or flaw in the system’s fusion engine. However, it rarely indicated anything about the quality of the detected attacks. Extraneous evidence didn’t necessarily have any correlation to lower or higher detection rates. Evidence Recall was even less useful. The thought was that as more evidence was used, the attack detection would be more accurate (higher recall and precision). However, empirically we found almost no relationship between the amount of evidence used and the quality of the detections. In fact, it almost appeared that the less evidence used the better the detections. This almost indicates that there are only a few truly relevant network events that indicate attacks. An open area of research or question is whether there’s a minimally complete set of data or events that could indicate the presence of an attack.

2.5.3 Cost Utility

Cost utility is defined as the ability of a system to identify the “important or key” attack tracks with respect to the concept of cost. In [15], two cost utility metrics were described. Since that paper was written, the *weighted cost* metric as applied to the cyber domain is no longer in use. The intent of the metric was to capture or gauge the usefulness of the system by considering the types of attacks detected with a positive weight and penalizing the system for false positives with a negative weight. Different weights were also assigned to different categories of attacks. Weighted Cost is then a simple sum of the values assigned to the types of attacks detected including false positives divided by the sum of the values of the attack tracks in ground truth. Observations when using the weighted cost showed that it didn’t add any value in measuring the performance of a Cyber SA system. The metric that appears to have great value in measuring the performance of a Cyber SA system is the attack score. The attack score is an earlier version of the “activities of interest (AOI) score” described in [8]. Attack score is calculated as shown in equation 2.8.

$$\text{AttackScore} = \frac{\#of\ attacks\ in\ GT * \#of\ tracks\ in\ GT - \sum\ of\ positions\ of\ detected\ attacks}{\#of\ attacks\ in\ GT * \#of\ tracks\ in\ GT - \sum\ of\ \#attacks\ in\ GT} \quad (2.8)$$

Attack score tries to measure the presentation of a prioritized list of hypothesized attacks by counting how many actual attacks occur and how close their priority is to the top of the list. In effect, the attack score measures the ability of a system to perform situation assessment as defined in 2.3. The attack score is considered a cost utility measure because the lower in a prioritized list that the actual attack appears

implies that more work is performed before the actual attack would be considered or could be acted on. For example, if the actual attack was at the top of the prioritized list, a user's attention would be drawn to it first and action taken. If the actual attack appeared at position 25, then the user would have to consider or "look into" 24 other attack tracks before getting to the actual attack track that was important or required action. Analyzing the 25 tracks indicate a cost in time before being able to take an appropriate action. Thus, an ideal attack score, with all actual attacks (true positives) at the top of the list, would be 1.0. Anything less than an attack score of 1.0 means that some level of effort is expended considering incomplete tracks or false positives. Another way of looking at the attack score is that it measures a system's capability to assess the situation. How the attacks are listed (prioritized) is determined algorithmically and could be influenced by the desire of the user to include; most critical, most damaging, most likely, greatest mission impact, etc. The different prioritization algorithms could influence the ordering of the list which in turn affects the attack score which could also provide differing assessments of the situation. How well this measure can assess the situation and what other metrics may be needed is a potential area of future research. The attack score has tremendous opportunity to be an indicator of improved analysis as more sensor types are considered, as better models are used in the fusion process, or as we improve our capability to add additional methods for situation analysis as described in [8]. With improving analysis, the attack score would approach the ideal of 1.0.

2.5.4 Timeliness

The final dimension, timeliness is the ability of the system to respond within the time requirements of a particular domain. More specifically, timeliness would need to measure the time elapsed before a decision could be made or action taken. Because we're interested in awareness, it's not simply the time it takes to present or detect an attack or activity but also includes the time it takes to enable awareness of the activity so that a user could make a decision. Timeliness touches the border between a measure of performance and a measure of effectiveness. So far, timeliness is an area of future research. To conclude this section on measures of performance, it's important to note that a single metric is probably inadequate to characterize the performance of a system. Rather, a set of metrics measuring the various dimensions of the problem are needed to fully characterize and provide insight into the performance of the systems.

2.5.5 Measures of Effectiveness

By measures of effectiveness (MoE), we mean measures that consider how effective a system is in enabling a decision maker's situation awareness of the environment

of interest. Have the tools or techniques developed to support SA improved or hindered the decision making process? Are alternatives that may not be immediately apparent considered? Can a novice decision maker more quickly become an expert or at least make expert-like decisions? To date, minimal research has gone in to measures of effectiveness but we expect to begin researching MoE both in general and specifically for the cyber domain very soon.

2.6 Conclusion

This chapter has described a generalized approach to enabling situation awareness and how that approach can be applied to the cyber domain. This application then enables Cyber Situation Awareness and ultimately Universal Situation Awareness. The challenging key to this approach is in identifying the activities individual decision makers are interested in and need to maintain awareness of over time. Then once the activity of interests are identified and modeled, the observations necessary to identify the activity need to be defined. Evaluating the effectiveness of the capabilities supporting SA depends on cognitive processes and determining if the technique has improved decision making. Measuring effectiveness for this approach is an open area of research.

Acknowledgements The authors thank Mr. Mike Hinman, AFRL/RIEA; Dr. Moises Sudit and Dr. Adam Stotz, University of Buffalo; Dr. Shanchieh 'Jay' Yang, Rochester Institute of Technology; Mr. Jared Holsopple, Rochester Institute of Technology; and countless others for their valuable insights and contributions to this research. This chapter is approved for public release, case number 88ABW-2009-1866.

References

1. <http://www.wikipedia.org>.
2. D. S. Alberts, J. J. Garstka, R. E. Hayes, and D. A. Signori. Understanding information age warfare. In *DoD Command and Control Research Program Publication Series*, 2001.
3. J. Antonik. Decision management. In *Military Communications Conference 2007 (MILCOM '07)*, pages 1–5, Orlando, FL, USA, October 2007. IEEE.
4. E. Bosse, J. Roy, and S. Wark. Concepts, models, and tools for information fusion. In *ISIF*, page 43. Artech House, Inc, 2007.
5. M. Endsley. Toward a theory of situation awareness in dynamic systems. In *Human Factors Journal*, volume 37(1), pages 32–64, March 1995.
6. B. McGuinness and J. L. Foy. A subjective measure of SA: The crew awareness rating scal (cars). In *Proceedings of the first human performance, situation awareness, and automation conference*, Savannah, Georgia, USA, October 2000.
7. U.S. Department of Defense, Data Fusion Subpanel for the Joint Directors of Laboratories, and Technical Panel for C3. Data fusion lexicon. 1991.
8. J. Salerno. Measuring situation assessment performance through the activities of interest score. In *Proceedings of the 11th International Conference on Information Fusion*, Cologne GE, June 30 - July 3 2008.

9. J. Salerno, M. Hinman, and D. Boulware. Evaluating algorithmic techniques in supporting situation awareness. In *Proceedings of the Defense and Security Conference*, Orlando, FL, USA, March 2005.
10. J. Salerno, M. Hinman, and D. Boulware. A situation awareness model applied to multiple domains. In *Proceedings of the Defense and Security Conference*, Orlando, FL, USA, March 2005.
11. J. Salerno, G. Tadda, D. Boulware, M. Hinman, and S. Gorton. Achieving situation awareness in a cyber environment. In *Proc of the Situation Management Workshop of MILCOM 2005*, Atlantic City, NJ, USA, October 2005.
12. L. Snidaro, M. Belluz, and G. Foresti. Domain knowledge for security applications. In *ISIF*, 2007.
13. A. Steinberg, C. Bowman, and F. White. Revisions to the JDL data fusion model. In *Joint NATO/IRIS Conference*, Quebec, Canada, October 1998.
14. G. Tadda. Measuring performance of cyber situation awareness systems. In *Proceedings of the 11th International Conference on Information Fusion*, Cologne GE, June 30 - July 3 2008.
15. G. Tadda and et al. Realizing situation awareness within a cyber environment. In *Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications 2006*, edited by Belur V. Dasarathy, Proceedings of SPIE Vol. 624 (SPIE, Bellingham, WA, 2006) 624204, Kissimmee FL, April 2006.

Cyber Situational Awareness

Issues and Research

Jajodia, S.; Liu, P.; Cohan, V.; Wang, C. (Eds.)

2010, XII, 252 p. 20 illus., Hardcover

ISBN: 978-1-4419-0139-2