# Visualization for Rapid Assessment and Projection in Cyber Situational Awareness

**Robert F. Erbacher**
Department of Computer Science, UMC 4205
Utah State University
Logan, UT 84322
Robert.Erbacher@usu.edu

**Abstract** - *The goal of this research was to create a set of next generation cyber situational awareness capabilities with applications to other domains in the long term. The goal is to improve the decision making process such that decision makers can choose better actions. This exemplified our entire research process and was the reason for our extensive work on vulnerability and impact assessment. These techniques will allow analysts to make better decisions and better prioritizations than is currently being done. Thus, we put extensive effort into ensuring we had feedback from network analysts and managers and understood what their needs truly were. Vulnerability assessment provides a numerical score indicative of how likely an attack on a given system is to succeed. Impact assessment provides a numerical score indicative of how substantially an attack will impact a system. The goal in the long term is to determine early in the life cycle of an attack as to the nature of the attack such that its impact can be assessed and analysts can then accurately prioritize it. While we did extensive work on generating these numerical scores, our primary focus was on generating novel visualization designs to present this information succinctly to the analyst.*

**Keywords:** network management, cyber situational awareness, visualization, attack detection, attack resolution.

## 1. Introduction

Network managers and analysts currently have significant difficulties in keeping up with the number of malicious or malicious appearing events occurring on their networks. This is exacerbated by the scale of many of today's networks. With non-routable IP addresses, network managers can be dealing with tens of thousands of systems,. These systems range from printers and routers to typical desktops, laptops, and the ever increasing number of mobile devices such as iphones, portable game devices, kindles, etc. Currently, network analysts must acquire the set of malicious events currently occurring on the network, manually prioritize the events based on likely impact or some other internal criteria, analyze the events to identify their nature, and ultimately resolve the events.

This is a very time consuming process that visualization techniques are currently attempting to try and resolve. This process is further complicated due to the difficulty in determining which events are truly malicious and the fact that so many events are generally occurring at any one time on the network. Along these lines, current research has focused on the display of higher level events or attributes, rather than the low level network events. For instance, the representation of attack streams made up of the individual network events greatly reduces the problem scale and provides far more relevant and meaningful data. In our case, we focus on the representation of impact and vulnerability assessment scores as our primary attributes.

## 2. Background

This research presents a next generation visualization designs for cyber event situational awareness. The goal of these techniques is to provide network managers and analysts with more critical and pertinent displays of relevant data. Intrinsically, the goal is to go beyond the current situational awareness capabilities of current techniques. Situational awareness is the creation of abstract higher-level representations of the underlying raw data. It has less of a focus on analysis and more of a focus on immediate comprehension. Situational awareness is:

> "knowing what is going on so you can figure out what to do" [1]

For situational awareness, we used Endsley's model [3] which essentially contains 3 stages of situational awareness. This includes perception, comprehension, and projection. The goal with situational awareness is to answer select questions rapidly:

- What is happening?
- Why is it happening?

- What will happen next?
- What can I do about it?

Thus, for our work in cyber situational awareness, the goal is to represent characteristics of attacks and attack events in a succinct form such that these questions can be answered. By rapidly providing analysts with this feedback as to the nature of attacks, based on identified events, the analyst should be able to make better and faster decisions. To achieve this end, we focused less on the individual attack events but rather focused on derived values, namely the impact and vulnerability assessments mentioned previously.

Prior research has focused primarily on level 1 of for comprehension through impact and vulnerability assessment, which amount Endsley's situational awareness model, namely providing for perception of events. This is exemplified in [8]. We focused this work on both levels 1 and 2 of Endsley's model, adding extensive capabilities to prioritization techniques for the events. Additionally, we provide examples of how level 3 in Endsley's model, projection, can be supported in the new visualization designs. The goal is to improve the decision making process such that better actions are taken. This exemplified our entire research process and was the reason for our extensive work on vulnerability and impact assessment. These techniques will allow analysts to make better decisions and better prioritizations than is currently being done. Thus, we put extensive effort into ensuring we had feedback from network analysts and managers and understood what their needs truly were.

## 2.1 Visualization

Data Visualization is the creation of graphical images for the representation of data through either abstract or physical relationships [9]. Such data visualization has garnered enormous interest in recent years due to the need for exploration and analysis of enormous volumes of data and the inability of automated techniques to provide the needed analysis. Visualization also maintains the user in the loop, allowing for intuition and expertise to take part in the analysis process. For example, bioinformatics-based data sets can be huge, in the GB range, with hundreds or thousands of parameters. The problem with such data sets is that it is not known what the analyst should be looking for in such a data set. Thus, automated techniques are limited in their effectiveness. The human analyst must direct the exploration and analysis process. Relying on solely textual responses from the analysis tools is a slow process as the human analysts must interpret and correlate the huge amounts of raw data with large amounts of computed data.

Thus, visualization attempts to represent both data in its raw form as well as the computed values, i.e., often statistical or mathematical results. In addition, when dealing with such large data sets the environment must provide for sampling, selection, or filtering of the data to ensure the most relevant information is presented on the screen. It is unfeasible to assume that even future display capabilities will be able to visually display the volume of data being discussed; especially considering the rate of growth of data collection processes in comparison with the growth rate of display technology.

The success of visualization derives from its reliance on human perception. Humans are able to visually interpret enormous amount of information that through other forms would be extremely slow and tedious. More specifically, reading textual information is considered a perceptually serial process as the reader must perceptually interpret each character in sequence to interpret a word and subsequently each sentence. A graphical image on the other hand can be interpreted in parallel, allowing a conceptualization of the image to be interpreted essentially instantaneously. There are particular visual characteristics that humans are particularly noteworthy at identifying and interpreting. These are traditionally termed pre-attentive components [4]. Such components include: texture, color, size, orientation, etc. These components allow the identification of anomalies, trends, and similarities essentially instantaneously. Thus, perception is a critical component of visualization research.

## 2.2 Vulnerability Assessment

The goal with vulnerability assessment is to identify the relative susceptibility of a system to attack. A system with a low vulnerability score is less likely to have a successful attack against it. Thus, an analyst will be more concerned with attacks against systems deemed to be more vulnerable. This will be further impacted by the priority of the system itself; i.e., analysts will be more concerned with attacks against a server than a common desktop.

We had two specific goals with respect to vulnerability assessment. First, was the generation of actual vulnerability assessment data. For this, we focused on the application of the Common Vulnerability Scoring System (CVSS) [6][5]. Second, was the creation of visualization techniques to represent the vulnerability scores. This paper focuses on two visualization designs, one of which is designed to intrinsically support the representations of vulnerability assessment scores.

## 2.3 Impact Assessment

The goal with impact assessment is to provide a prioritization score to the analyst. In essence, analysts must

deal with enormous numbers of events on a daily basis, especially for larger organizations. These events must be prioritized in order to identify a sequence in which they will be resolved. Currently, this prioritization is adhoc and based solely on the analysts expectations of the event significance. The goal here is to create an impact assessment that identifies for the analyst the current and predicted impact of each event according to analyst specified metrics.

When computing an impact assessment score, in essence we are attempting to compute the amount of degradation of available resources an event or sequence of events is likely to impose on the network. More importantly, however, is identifying the levels at which the degradation in available resources negatively impacts our ability to use those resources. This can be thought of as cyber readiness; i.e., how ready is the network to deploy missions or countermeasures. One of the visualizations discussed in this paper is specifically designed to integrate and represent the computed impact factors to analysts.

## 3. Visualization Design Methodology

The visualization design process involved several steps. The goal was to use a human in the loop design strategy that ensured the resultant techniques were of value to the target user base. The design process had eight primary steps:

1. An initial brainstorming meeting was held with analysts, network managers, security researchers, and visualization researchers at PNNL. This resulted in an initial series of questions such as the time frame we needed to consider for this level of analysis. Such questions were primarily answered from existing Cognitive Task Analysis (CTAs), especially Anita D'Amico's from Secure Decisions [2] and Stefano Foresti's from the University of Utah [7].

2. This was followed by a set of individual interviews with network analysts and managers. The main impact of the individual interviews was the identification of needed future work, i.e., the need for a communication board. This was out of scope for this task but such capabilities are clarified and more fully detailed in later stages of the research.

3. Examination of previous work provided a third leg of background information, ideas, and concepts needed for the design. This resulted in a separate presentation documenting initial ideas and needs. The idea of overlays used in our designs was derived from background literature on visualization.

4. Next, a set of scenarios were generated to guide further discussions with network managers and analysts. These designs were directly derived from the initial background acquired during the previous three stages.

5. A small focus group was held with network analysts and managers to review the scenarios and acquire feedback and comments. This resulted in the development of a new task flow for network managers that led to the identification of the need for two levels of visualization techniques. In essence, one set of capabilities is needed for immediate assessment of network status and impact while a second set is needed for more detailed analysis of the network.

6. An initial set of visualization designs were created based on the acquired background information and analyst feedback.

7. Individual interviews with network managers and analysts were then conducted to acquire feedback on specific aspects of the designs.

8. Updated designs were created based on feedback from this last set of interviews

## 4. Rapid Assessment and Prediction Visualization Design

Figure 1 exemplifies the design of a visualization technique designed to meet the immediate needs of network managers and analysts. It is designed around multiple overlapping histograms, using transparency to maintain visibility of the individual histograms. This design includes two additional unique characteristics. First, is the ability to select one or more events in time. This allows the event to be tracked or for the event to be examined in more detail in a separate, more detailed, visualization. Second, is the ability for the environment to represent prediction. Determination of actual prediction was beyond the scope of this project, though we did determine how prediction could potentially be determined intelligently (at a high level). For the purposes of this project, we employed simple linear projection, in essence linearly projecting the current trends into future time points in order to determine the network status should current events not be resolved. Finally, the visualization design was created to incorporate interaction directly. Thus, the data impact parameters to be represented can be directly selected. Additionally, the selection of events can be done through direct manipulation. Of critical importance is the fact that analysts will need an additional visualization for more detailed analysis, one such design is presented in the next section. Again direct manipulation is incorporated into the

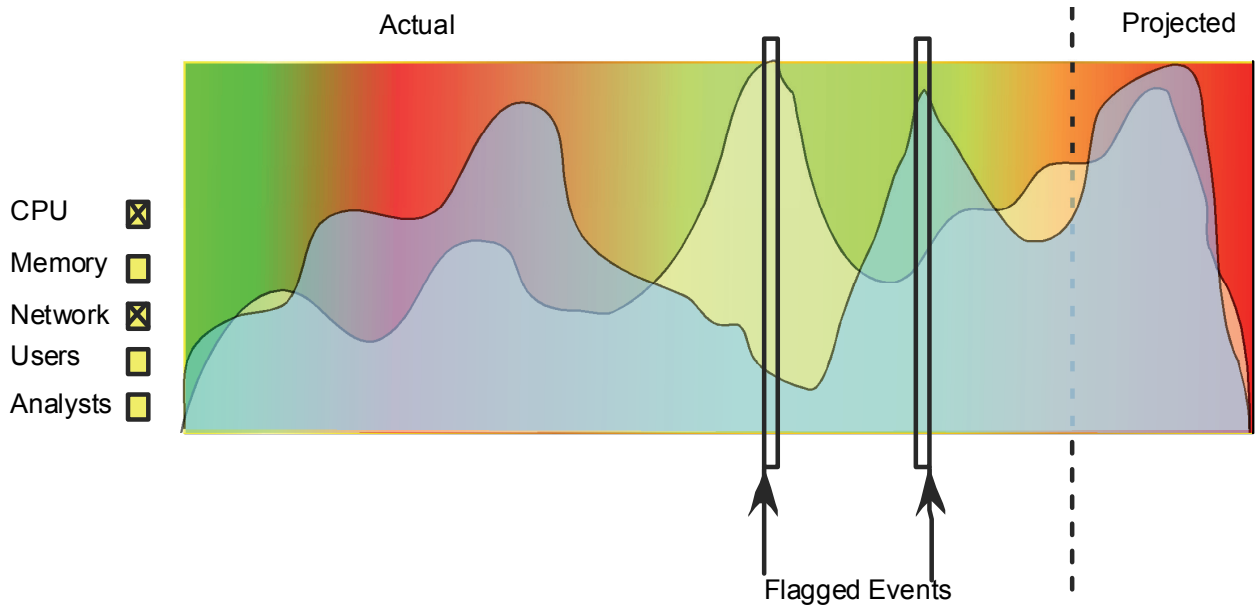


**Figure 1:** Rapid assessment and prediction visualization design concept.

design to allow for the passing of specific event conditions and periods of time to the detailed analysis visualization design.

A more intelligent projection can be computed by correlating known vulnerabilities on a given system with identified attacks occurring elsewhere on the network. This can then be mapped to the known impact of the identified attacks. This will provide a good prediction of the expected impact of an attack on individual systems and the network as a whole. In the end, this could then be extended to missions.

This visualization design is specifically geared towards meeting the needs of phase 2 and phase 3 of Endsley's situational awareness model. It does this through the representation of both high level network events as well as computed and projected impact assessments. This immediately identifies to the network manager the fact that a malicious event has occurred and the likely impact the event will have should it not be resolved. This design goes to the heart of situational awareness as it conveys true comprehension and meaning to the network manager. This goes well beyond the typical but limited perception capabilities provided by existing situational awareness techniques. Finally, the projection capabilities of this design meet the needs of the third phase of Endsley's model.

## 5. Detailed Analysis Visualization Design

Figure 2 provides the detailed visualization design concept in which additional content is embedded into the network topology itself. This visualization design works directly with the visualization design from the previous section to allow for detailed analysis of an event and determine the exact nature and resolution patch for an event. The visualization design does attempt to maintain aspects of situational awareness. In essence, the visualization design provides for perception of the current network status. Acquiring comprehension requires more exploration within the environment than is required by the first visualization design concept.

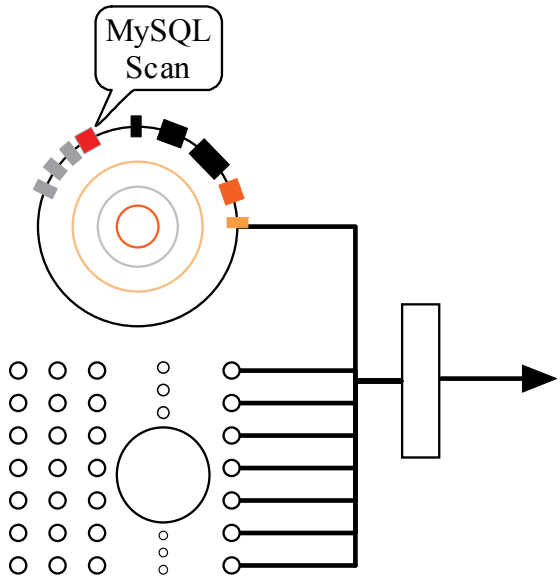


**Figure 2:** Detailed analysis visualization design concept.

In this visualization design, each node in the network topology is automatically scaled by the number of events that node is receiving. Characteristics of those events are then

incorporated into the ring of the node. For instance, each type of event can be given a different location around the ring and a different color. A hash mark on a given position around the ring indicated the presence of that event. The thickness of the hash mark represents the number of events of a given type that were received. Multiple rings can be used to show older periods of time, providing an indication of changes over time.

Given the identification of an event from the previous visualization design, the analyst can then examine this visualization design, immediately identify which systems are likely involved, and identify what events preceded and have followed the identified malicious event. This will provide for the ability to rapidly identify malicious events followed by the ability to rapidly assess, analyze, and resolve said events.

## 6. Future Work

Feedback from network analysts and managers has so far been positive with respect to the current visualization designs. However, full implementations of the visualization designs need to be completed to allow for complete user studies to be performed. Additionally, while we identified many of the needed interaction techniques, actual research is needed to develop novel interaction techniques to more substantially improve upon the effectiveness of the visualization techniques.

Additionally, we have not dealt with the issue of collaborative analysis or mobile devices. With the proliferation of mobile devices, the visualization techniques must be adapted to work in such environments. In particular, this will require further work with interaction techniques. Finally, the visualization techniques must allow analysts to work with one another, even when some of the analysts may not be local. This becomes critical with the need to use less experiences network analysts. Should such a novice analyst identify a suspicious event they could dynamically pass what they've learned to more senior analysts, likely to the senior analyst's mobile device, to acquire feedback as to a more experienced interpretation of the event.

## 7. Conclusions

We followed a human in the loop research process, creating a new task-flow model for network management that will greatly impact future research in this domain. Analyst interviews identified the basic requirements, critical parameters, and characteristics needed for the next generation of cyber situational awareness visualizations. These analyst interviews resulted in a new view of the needed capabilities for network analysts and managers.

This project clearly shows that impact and vulnerability assessment can be used by network analysts and managers to improve the analysis of network events. Additionally, we created effective next generation situational awareness visualization techniques for the display of traditional cyber data as well as the vulnerability and impact assessment data. Combined, these capabilities will go a long way towards improving the cyber management challenges being seen in today's network environments. The generated visualization designs met the two identified needs for situational awareness visualization, namely immediate summary representations and long-term analysis representations. An additional advantage of the summary representations is its direct solution for the scalability issues inherent in cyber situational awareness.

While the current mockups were designed for cyber situational awareness, they are in fact designed to be completely generalizable. This is particularly true of the summary visualizations. The main issue will be with the implementation and the need to support generalizable data in the long term. Thus, there is no limitation that the techniques only be applied to cyber data. They could easily be applied to physical assets of any type by any domain.

## 8. Acknowledgements

Many students played significant roles in the performance of the project, including: Anupama Biswas, Anusha Davuluri, Chris Harris, Srinidhi Kakani, Stephen Miller, Steena Montiero, Sarah Moody, Rian Shelley, and RB Whitaker. This research was funded in part by AFRL under project FA8750-07-C-0163.

## 9. References

1. Adam, E. C. (1993), "Fighter cockpits of the future," *Proceedings of 12th IEEE/AIAA Digital Avionics Systems Conference (DASC),* pp. 318-323.

2. Anita D'Amico, Daniel Tesone, Kirsten Whitley, Brianne O'Brien, Emilie Roth, "Understanding the Cyber Defender: A Cognitive Task Analysis of Information Assurance Analysts," Report No. CSA-CTA-1-1. Secure Decisions. Funded by ARDA and DOD.

3. Endsley, M. R. (1995b), "Toward a theory of situation awareness in dynamic systems," *Human Factors*, 37(1), pp. 32-64.

4. Gordon, S. E. and Gill, R. T., "Cognitive Task Analysis," In C. Zsambok and G. Klein (Eds.). *Naturalistic Decision Making*, Mahwah, NJ, Lawrence Erlbaum Associates, 1997, pp. 131-140.

5. Phister, P. W. Fayette, D. Krzysiak, E., “The CyberCraft Concept Linking NCW Principles with the Cyber Domain in an Urban Operational Environment,” *MILITARY TECHNOLOGY,* Vol. 31, No. 9, 2007, pp. 123-131.

6. Mike Schiffman, Gerhard Eschelbeck, David Ahmad, Andrew Wright, Sasha Romanosky, "CVSS: A Common Vulnerability Scoring System", National Infrastructure Advisory Council (NIAC), 2004.

7. Stefano Foresti and Jim Agutter, “Cognitive Task Analysis Report,” University of Utah, CROMDI.. Funded by ARDA and DOD.

8. Stefano Foresti, James Agutter, Yarden Livnat, Robert Erbacher, and Shaun Moon, "Visual Correlation of Network Alerts," *Computer Graphics and Applications*, Vol. 26, No. 2, March/April 2006, pp. 48-59.

9. Wolfe, J.M., “Visual Attention,” *In Seeing: Handbook of Perception and Cognition*, K.K. De Valois (Editor), Second Edition, Academic Press, pp. 335-386, 2000.