CyberSAVe – Situational Awareness Visualization for Cyber Security of Smart Grid Systems

William J. Matuszak
Adaptive Methods
Two Corporate Pl, Suite 203
Middletown, RI 02842
(401) 380-2080
bmatuszak@adaptivemethods.com

Lisa DiPippo
University of Rhode Island
Kingston, RI 02881
(401) 874-2701
dipippo@cs.uri.edu

Yan Lindsay Sun
University of Rhode Island
Kingston, RI 02881
(401) 874-5803
yansun@ele.uri.edu

ABSTRACT
We offer algorithms and visualization techniques for cyber trust in a Smart Grid system. Cyber trust is evaluated in terms of a mathematical model consisting of availability, detection and false alarm trust values, as well as a model of predictability. We develop a prototype Cyber Situational Awareness Visualization (CyberSAVe) tool to visualize cyber trust. We provide Operational Decision Aids (ODAs) displayed in context with a SCADA management information. We define cyber trust metrics, which are calculated and displayed in real-time in the Metric Assessment System (MAS) of CyberSAVe. We demonstrate the use of trust combined with visualization of trust to detect various types of attacks on the Smart Grid.

Categories and Subject Descriptors
K.6.m [**Management of Computing and Information Systems**]: Miscellaneous - Security

General Terms
Algorithm, Performance, Design, Experimentation, Security, Human Factors.

Keywords
Cyber Security, Trust, Visualization, Situational Awareness. Framework, Smart Grid.

1. INTRODUCTION
SCADA (Supervisory Control And Data Acquisition) systems add a new potential for damage that can be inflicted by hackers and cyber terrorists. Using complex attacks, hackers can seize control of and inflict damage using the physical system being controlled by a SCADA. SCADA system managers must maintain situational awareness of their networks, sensors, and controllers to avoid and/or minimize such attacks. In this paper we will use an electrical Smart Grid as an exemplary SCADA system. A *Smart Grid* has added vulnerabilities as it is operated in the public domain and has many connection points.

The sensors in a Smart Grid system can be as vulnerable to attack as the actuators and the controllers. Thus, it is crucial that a secure Smart Grid is able to determine if sensors are behaving correctly and non-maliciously, and to act on this information. In a Smart Grid system, a sensor node, like a phasor measurement unit (PMU), is designed to detect and report fluctuations in power that is being transmitted. If such a sensor is compromised, it might send false alarms to nearby actuators, it might not report existing fluctuations, or it might stop reporting altogether. In any of these cases, the Smart Grid system will be vulnerable to attacks on the transmission lines, the substations, or on the generating station since the information being monitored by the sensor cannot be trusted.

Trust is an essential component to any system in which nodes collaborate and where humans monitor and act on the observed behavior of the network. In a Smart Grid system, the information gathered from measurements and sensors must be trustworthy if sensor feedback is to be relied upon to control system components. Similarly, directions from controllers and actuators must be trustworthy if controller commands are going to alter the state of a physical system. If any node is compromised, it can no longer be trusted. This trust information should be part of the control loop of a Smart Grid system. Further, it is important that this trust information be available to operators to provide accurate situational awareness of the Smart Grid system.

To address these issues, we have developed CyberSAVe (Cyber Situational Awareness for Visualization), which provides visualization of trust in the nodes of a network in order to allow a user to make critical decisions. CyberSAVe incorporates our mathematical model of cyber trust that includes aggregation of various types of trust, as well as a notion of trust redemption. It provides a set of visual tools that an operator can use to investigate and explore a situation in order to, not only see where problems arise in the network, but also to understand the causes of the problems, and how to mitigate them.

In Section 2 of this paper, we present our trust model, which we have incorporated into CyberSAVe. In Section 3 we describe visualization techniques that we have developed to allow human operators to understand the trust situation within the context of a Smart Grid system. In order to demonstrate the effectiveness of our trust-based visualization, we implemented a series of tests in the simulated Smart Grid environment. Section 4 describes these tests. We present related work in Section 5 to demonstrate the novelty of CyberSAVe and show how it improves upon existing research. Finally, we conclude in Section 6 with a discussion of the impact of this work and future applications of trust visualization.

VizSec '13, October 14 2013, Atlanta, GA, USA
Copyright 2013 ACM 978-1-4503-2173-0/13/10…$15.00.
http://dx.doi.org/10.1145/2517957.2517961

2. CYBER TRUST

Trust is an essential component to any collaborative relationship. For instance, in a Smart Grid system, the information gathered from sensors must be trusted in order for self-healing algorithms or humans in the loop to act on it. In previous work, we developed a theoretical foundation for trust evaluation in wireless sensor networks [5,6]. We have extended this work to apply to the behavior and expected threats in a Smart Grid system. We use a Smart Grid system as an illustrative application, but the fundamental concepts that we develop can extend to any SCADA system in which sensor nodes may behave maliciously.

The trust model that we have developed has three essential components: trust computation, multi-dimensional trust metrics and mathematical properties. We describe each of these components in the following subsections.

2.1 Trust Computation.

A trust relationship is established between two parties for a specific action. We introduce the notation: *{subject*: *agent; action}* to represent that the subject trusts the agent to perform the action. So for example, in a Smart Grid, an actuator may trust a PMU sensor to detect and accurately report on a power fluctuation *{actuator: sensor; report_fluctuation}*. The trust, *T*, between these components is computed based on observed behaviors using the following Bayesian-based formula:

$$T = \frac{GB + 1}{GB + BB + 2} \quad (1)$$

where *GB* represents a good behavior, such as a reported fluctuation, and *BB* represents a bad behavior, such as a missed report of an existing fluctuation.

2.2 Multi-dimensional Trust Metrics

Trust in the nodes of a Smart Grid system can be multi-dimensional. That is, it can be defined by more than one behavior that the node performs, and therefore, it may require aggregation of several types of trust value. A PMU sensor node is designed to detect and report fluctuations in power that is being transmitted. Thus, we compute the following types of trust based on the behaviors of these nodes: *detection trust* represents how much we trust a sensor to detect and report a power fluctuation; *false alarm trust* represents how much we trust a sensor to report only when a power fluctuation is detected (no false alarms); *availability trust* represents how much we trust a node to respond to a request for a report. To aggregate these types of trust, we compute *overall trust* (*OT*) as a weighted sum as follows:

$$OT = \sum_{k=1}^{n} T_k * W_k \,, \sum_{k=1}^{n} W_k = 1 \quad (2)$$

where T_k is the trust value for the k^{th} type of trust, and W_k is the weight that is placed on the k^{th} type of trust. Overall trust can be used by the Smart Grid system to make decisions about how to self-heal, or it might be used by a human operator to decide to dispatch repair units when an event is detected.

2.3 Mathematical Properties of Trust

Trust is established based on observed behaviors of an entity. In order to quantitatively evaluate trust we identify mathematical properties of trust values. *Direct trust* is established when the behavior of node A is directly observed. For instance, if node A reports to a local programmable logic unit (PLC) that it has detected a power fluctuation at a particular time but other nodes reporting to the same PLC contradict this report, then the *detection trust* of node A may be reduced. *Trust initialization* affects how fast trust values converge. In our simulations of a Smart Grid system, we assumed an initial trusted deployment, so we set initial trust value of all nodes to 1, which means fully trusted.

Trust is a dynamic characteristic. A good node may be compromised and turned into a malicious one, while an incompetent entity may become competent due to environmental changes. In order to track these dynamics, an observation made a long time ago should not carry the same weight as one made recently. A trust redemption scheme allows a node's trust value to be redeemed with time and with subsequent good behaviors [5]. Such a scheme allows for a single bad behavior to be forgotten more quickly than multiple bad behaviors.

Trust redemption is useful to allow for better utilization of system resources when nodes behave badly for non-malicious reasons, e.g. environmental anomalies that cause sensors to malfunction. However, a sophisticated attacker may understand that these redemption algorithms are being implemented and try to fool the trust system into allowing repeated bad behaviors as long as there are enough good behaviors or time to allow trust to be recovered. This type of attack is known as an On/Off attack because it turns the attack on and off in order to fool the trust system. For example, if a PMU sensor sends a false alarm, the trust of that node would be reduced. If, however, the node then performs correctly for some time, its trust will be redeemed. Repeating this behavior will allow the node to continue to cause problems in the system and not be detected.

We have developed a new type of trust, called *predictability trust* [5], that is used to modify the overall trust of a node based on how well the current trust value predicts future behavior. We have shown that using predictability trust to compute the overall trust can detect an On/Off attack. Further, we have shown that a system designer can set parameters to choose the attack ratio that can be detected, traded off with system resources used [5].

Predictability trust is computed based on how well a node's behavior meets expectations. For example, suppose a node's current detection trust is 0.9, meaning that we (or the system) predict this node will detect at least 90 % of the power fluctuations in its range. If this node detects above 90% of the fluctuations, it meets the prediction, and is considered to perform a good predicted behavior (GPB). If this node detects below 90% of the fluctuations, it does not meet the expectation, and thus performs a bad predicted behavior (BPB). We count the number of GPBs conducted by node *i*, denoted by GPB_i , and the number of *BPBs* conducted by node *i*, denoted by BPB_i. The *predictability trust* of node *i* is computed as follows:

$$PT_i = \frac{GPB_i + 1}{GPB_i + BPB_i + 2} \quad (3)$$

Predictability trust is used in two ways in our trust computation. First, it is used to set the maximum value of the overall trust using the following formula:

$$OT_{new} = OT_{old} * PT \quad (4)$$

The overall trust that is computed with the weighted sum in Equation (2) is updated by multiplying the predictability trust. This way, decisions made based on the overall trust will take into account how predictably the node has behaved. Second, predictability trust is used to control the speed of trust redemption using the following formulas:

$$RF = (OT_{new} * PT) * \alpha + 1.0 \quad (5)$$

$$T_k = T_k * RF,\ k=1 \ldots n \quad (6)$$

RF is the redemption factor computed using overall trust and predictability trust, and $0<\alpha<1$ represents a tolerance setting where a lower value of α provides for a stricter security tolerance. Then each of the trust values (T_k) is multiplied by the redemption factor *RF* to allow it to be redeemed periodically with time. Thus, by incorporating predictability trust into the computation of overall trust, the visualization of trust that we present in the next section can alert the human operator of the erratic behavior of a sensor node.

3. CyberSAVe - VISUALIZATION FOR SITUATIONAL AWARENESS

Understanding the level of trust in specific nodes of a Smart Grid system is crucial to isolating the attack and removing nodes from the system in response. Visualization of certain characteristics of nodes in the system can provide valuable insight into the nature of an attack. In our CyberSAVe system, we have implemented data aggregation algorithms to allow an operator to choose characteristics on which to focus. CyberSAVe computes various other metrics that assist a human operator in better understanding a critical situation. Further, CyberSAVe provides operational decision aids that allow for straightforward visualization of these metrics and aggregation techniques. In this section, we first describe the data aggregation that we have implemented in CyberSAVe. We then discuss the metrics that CyberSAVe can display in various types of visualization, and the operational decision aids implemented in the system. The section goes on to present the visualization framework and overall architecture of the system. This section concludes with a demonstration of some of the graphical aspects of the system.

3.1 Data Aggregation

In monitoring the security of a Smart Grid system, it may be important to consider trust data of nodes that share certain common characteristics. In CyberSAVe, we have implemented a mechanism that facilitates the visualization of trust aggregated over specified parameters. This type of aggregation can provide awareness of potential risks and threats that may not be obvious otherwise. The parameters that we list below were chosen to represent common characteristics of Smart Grid nodes. In general, CyberSAVe is capable of aggregating over any defined parameters in the application domain.

- Geographical Location – Provides the capability to sort and analyze data by geographical location of the sensor, control system and/or the physical system. Aggregation of data by geographical location can highlight physical attacks and local jamming. As a defensive measure operators will be able to deactivate or increase vulnerability values of nearby sensors.
- Network Parameters – Sorts data based on the Internet Protocol (IP) address and subnet of the nodes in the Smart Grid system. Trust data aggregated by network parameters can alert operators of potential network intrusions and allow defensive actions such as isolating a subnet.
- Equipment Manufacturer – Groups data by the manufacturer of the Smart Grid sensors and controllers. Analyzing trust by sensor manufacture can help operators detect and identify complex worms and malware introduced in the production line of equipment. Detection of a cyber attack that is occurring on a specific piece of equipment at multiple sites will allow operators to isolate like equipment at other sites.
- Software (SW) Environment – Aggregates trust data based on the software environment (e.g. operating system, SW drivers, 3rd party commercial SW) used in components of the Smart Grid system. This can help detect malware associated with specific commercial software.
- Trust Type – Aggregates and displays data by trust type (availability, detection, false alarm discussed in 2.2). Trust type provides insight into type of attack. A complex attack that causes missed detections is likely from a highly funded cyber terrorist group or a nation state.

Figure 1 demonstrates several types of aggregation that CyberSAVe can display. We can see from the figure that while aggregating by voltage or by company does not show a significant difference in trust, when we aggregate by the senor type, one particular sensor has a much lower trust than the others. This can be a strong indication that that sensor has been compromised. An operator can use this information to decide to inform all substations using that sensor type to ignore the sensor data.

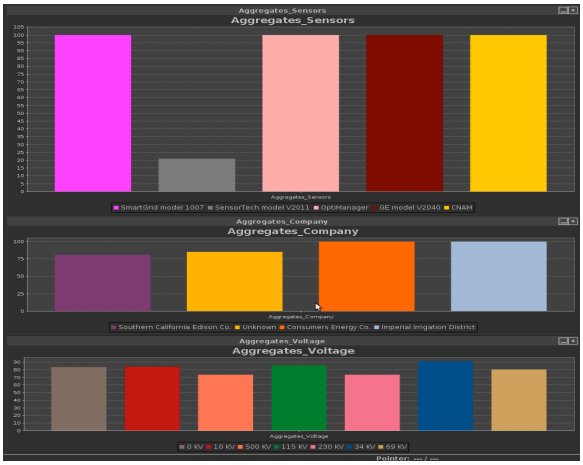


Figure 1: Aggregated Trust

3.2 Cyber Trust Metrics

Metrics provide quantized measurements of system performance. In CyberSAVe, we collect and display metrics related to cyber trust and system performance in real time. Our Metric Assessment System (MAS) provides the capability to analyze data in real-time via histogram, bar graph, or time history plot. The metric data is linked and displayed in context with an operational picture. CyberSAVe can display trust metrics for a selected Smart Grid system node or a group of nodes. MAS is configured to alert when a trust value falls below a defined threshold. The threshold and the time window (redemption factor) are configurable based on threat and environment considerations. Configurable thresholds allow an operator to adapt his posture to the current security environment, ensuring the system is not static. Operators can manually update or override system trust values by using controls.

Among the metrics that CyberSAVe can provide are:

- Single node or multi-node historical trust versus time
- Bar graph of trust by aggregated value
- Histogram distribution of trust for multiple sensors over time
- Percent of nodes with high / low trust over time
- Birds eye view of current (instantaneous) trust and on geo
- Individual trust values or overall trust values for nodes
- Aggregate trust values over geographic area (city, state, etc.)

As an example display for an individual node, Figure 2 shows a histogram of *availability trust*. A histogram can help an operator understand environmental conditions and/or threat posture. In Figure 2 we see that the mean availability trust is low and the

variance is relatively high. This can indicate a poor environment compared to a node with a high mean availability trust and a small variance.

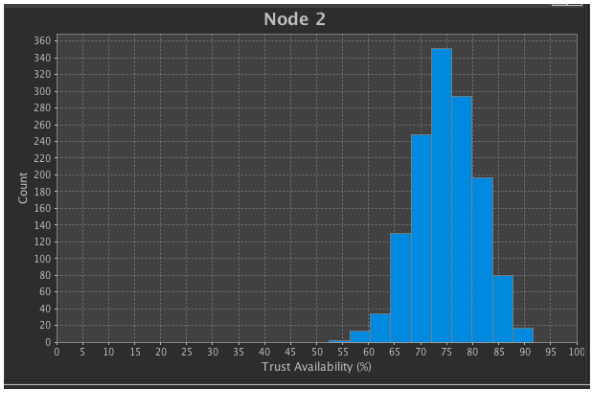


Figure 2 – Histogram of Availability Trust

Figure 3 illustrates the visualization of all three types of trust that we implemented in our Smart Grid simulation, plus the overall trust that is a composition of the other types of trust, along with predictability trust. From this figure we can see how the dips in false alarm trust cause the overall trust to also go down. The overall trust, shown in yellow at the bottom of the figure, redeems its value more slowly than the false alarm trust because of the incorporation of predictability trust into its calculation.

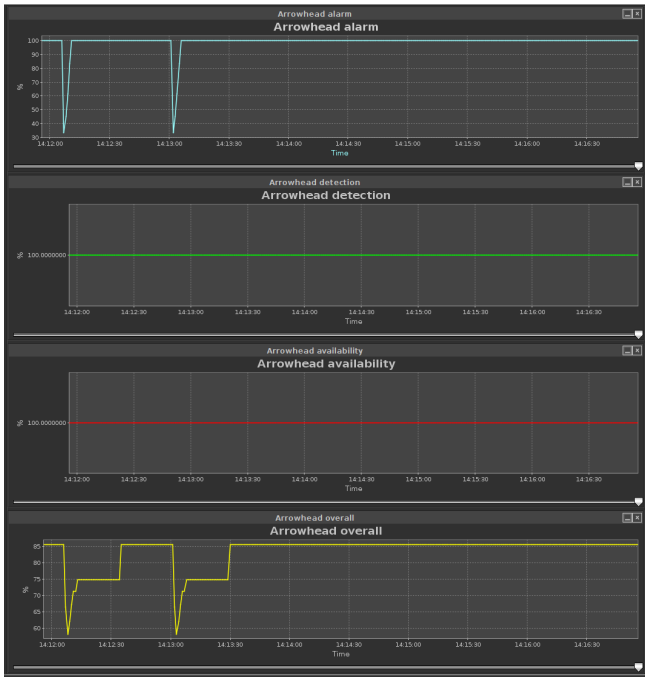


Figure 3: MAS display of cyber trust over time

3.3 Operational Decision Aids

CyberSAVe includes several visualization techniques and Operational Decision Aids (ODA) providing situational awareness of system assets. In order to display trust information in context with the overall operational domain, ODAs are provided as overlays on a reference model operational picture.

3.3.1 Operational Trust Indicator (OTI)

The Operational Trust Indicator (OTI) is a decision aid that displays all three types of trust into one indicator, using color to represent trust value. Figure 4 is an example of the OTI. Each of the three types of trust is represented in a section of the outer ring of the circle. In Figure 4 we see that *false alarm trust* is above the specified threshold, and so is depicted green. *Detection trust* is in a "suspicious" range, so it displays as yellow. And *availability trust* is below the warning threshold and is therefore red. The small circle in the center represents the *overall trust*, computed as a weighted sum of the other types of trust, and updated by *predictability trust*, as described in Section 2.3. Any node that is monitored in the Smart Grid System will have associated with it an OTI so that the trust can be visualized by the operator. The OTI for each node can be overlaid in either a geographical plot or a network schematic. Operators can drill down on the OTI and receive details of trust values and review associated metrics.

Trust Figure 4: OTI with Default Threshold Values

3.3.2 Contextual Trust Visualization

Cyber analysts and decision makers need cyber ODAs in context with their operational picture in order to make best use of the information. CyberSAVe does this by via overlays and Operational Decision Aids on a reference model operational picture. The ODAs and overlays can be transitioned directly to an operations management display. Figure 5 shows an example of a geographically dispersed Smart Grid system. Each "dot" represents a node geographically located in the system. The dots at this level are colored based on their overall cyber trust value. An operator can "drill down" on any node to view a time history and metrics of trust for the selected node. The image in Figure 5 shows the OTI for a specific node, along with instantaneous trust and overall trust over time for the node.

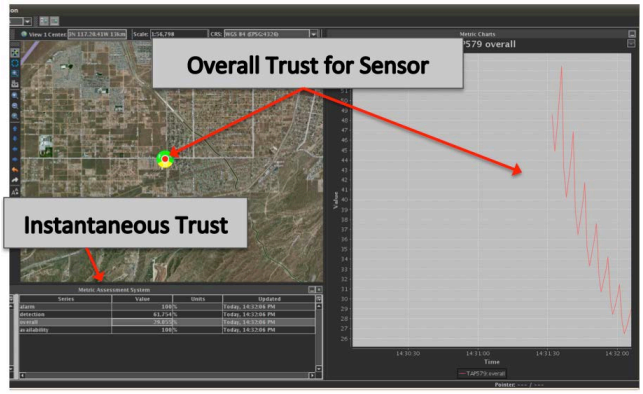
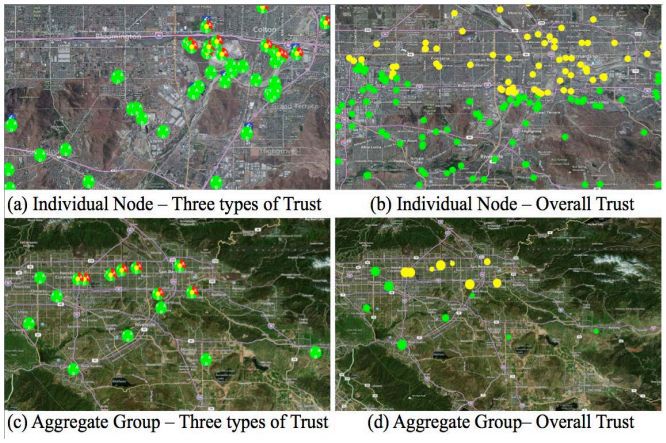


Figure 5: CyberSAVe display with ODAs and Metrics

3.3.3 Aggregated Trust Visualization

As we discussed in Section 3.1, CyberSAVe provides the ability to aggregate trust data across various characteristics. This aggregation can be overlaid on the contextual display to allow for visualization of trust based on these characteristics. Figure 6 demonstrates how aggregation can be visualized at various levels. CyberSAVe has operator controls to allow switching among these different views. Figure 6(a) shows the trust values for each individual node and each type of trust. This view may be too cluttered for an operator, so CyberSAVe has a zoom capability to

drill down in the image. Figure 6(b) is the same set of nodes, zoomed slightly, and only displaying the overall trust. From this view it is easier to discern problem nodes. Figure 6(c) depicts an aggregation of the nodes by city. This view combines the trust values of all nodes in the same city and displays this aggregation with all three types of trust in the OTI for the city. Finally, Figure 6(d) shows the aggregation over the cities on the map with overall trust displayed for each city. This view also uses the size of the dot to depict the number of nodes in the aggregation. The aggregate dots are located in the mean location on the map.



(a) Individual Node – Three types of Trust (b) Individual Node – Overall Trust

(c) Aggregate Group – Three types of Trust (d) Aggregate Group– Overall Trust

Figure 6 - Aggregated Trust Visualization

3.4 Visualization Framework

We provide a Visualization Framework (VF) as a deployment environment for the CyberSAVe prototype. Figure 7 depicts architectural concepts of our framework exhibiting several service oriented application models

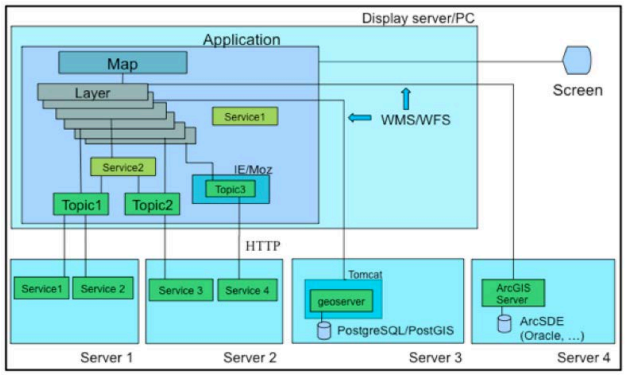


Figure 7 – Visualization Framework Architecture

Our Visualization Framework concept provides a flexible application framework for geospatial application development and deployment supporting open system specifications, open GIS toolkits and services, and open Application Programming Interfaces (APIs). It is ideal for the purpose of ascertaining and conveying trust level. Modules loaded into the framework use framework services to define geographically registered overlays (map overlays) and associated user interface components such as panels and toolbar buttons that implement the desired controls for the application layer. Each module may add its own overlays and controls without an explicit need to coordinate with any other module. This allows the application to be deployed with a mix of modules tailored to the deployed system. The module system permits module development teams to focus on providing unique capabilities without concern for any other module to be deployed concurrently. Development teams only need adhere to the common infrastructure. Supporting the module system are common infrastructure services that allow modules to interact with framework components, and with other modules cooperatively. Interactions between modules are managed using two key mechanisms – a *Service Registry* and an *Event Service*. In addition, the *Map Services* provide the ability to layer multiple levels of maps on the visualization. Since all tools and features for any application are loadable modules, platform-specific and application-specific functionality is integrated using the module loading system, and the composition of platform and application features is defined by runtime configurations/scripts.

3.4.1 Service Registry

The Service Registry allows modules to locate dynamically-registered services provided by the framework itself and by other modules. Operations supported by the Service Registry include service discovery, service change notification, and service enumeration. Modules may register an abstracted service interface exposing the precise set of operations to be used by other modules. This abstraction decouples the specification of a service from the implementation of the service, allowing alternate implementations.

3.4.2 Event Service

The Event Service allows the framework to notify interested clients of framework state changes such as the completion of application startup, or the beginning of application shutdown. Modules that provide common services may define their own events for distribution by the framework Event Service. A module may subscribe to any event at any time and the event publisher is not required to have explicit knowledge of any of the event subscribers. This mechanism allows modules to be deployed independently and optionally.

3.4.3 Map Services

The framework manages a map composed of an ordered list of layers, where each layer represents content that may be visible to the operator. Layers are rendered in order, forming a composite picture. A variety of layer types may be created and added to the map, corresponding to different content models and data source types (WMS raster imagery, WFS vector overlays, dynamic vector feature overlays, etc.) The map is visualized using one or more map views, where a map view represents a set of visible map layers, plus an observer (or camera) position and other state information. Each view is shown displayed in a rectangular sub-window. A single main-map view is always visible, but a number of secondary map views may be created using map services. The user interface allows the operator to create and manage secondary map views, in addition to allowing the operator to choose what layers are to render for a given map view.

4. SCENARIO TESTS

To demonstrate our implementation of trust visualization in CyberSAVe, we developed a set of test scenarios in which we simulated attacks of varying complexity. In this section, we describe those tests, discuss how our trust algorithms handle each one, and show how CyberSAVe visualization tools yield clear, understandable feedback to operators monitoring the system.

4.1 Scenarios

Possible attacks on a power distribution Smart Grid range from physically destroying equipment to advanced attacks at multiple sites. Here we discuss the application of cyber trust computation and visualization against varying levels of attacks, compared to

more traditional defenses. Figure 8 summarizes the types of attacks we simulated along with the types of defense that can work to protect against each. In the figure, a check in the box indicates that the method in the box, along with each previous method is sufficient to defend against the attack in the specific row.

Protection vs Attack	Sensor	+Firewall	+Cyber Trust	+Predictability	+Visualization
Physical Attack (PA)	✔	✔	✔	✔	✔
Simple Cyber Attack (SCA)		✔	✔	✔	✔
Intermediate Cyber Attack			✔	✔	✔
Advanced Cyber Attack				✔	✔
Advanced Collusion Attack					✔

Figure 8 – Scenario Tests

Physical Attack – A physical attack causes power failures by such means as cutting cables or dropping trees on power lines. Physical problems in a Smart Grid system are detected by traditional sensors, as long as the sensors are working as expected. Row 1 in Figure 8 shows that the presence of sensors in the network can defend against this type of attack.

Known Malware Attack – A computer virus installed on a sensor node in a Smart Grid system can be used to slow or disrupt operations. These attacks can be handled by traditional anti-virus software and firewalls. Row 2 of Figure 8 indicates that, since the sensor itself may be attacked, it is not defense enough to defend against a malware attack, and firewall technology is necessary along with the sensors.

Simple Cyber Attack – A simple cyber attack on a Smart Grid is an attack on the sensor nodes that evades firewalls by any of a variety of means including virus installed by an insider at manufacturing sites. This type of attack will cause the sensor node to send false alarms, miss detections of power fluctuations, or fail to report anything at all. Combined with a physical attack, this type of attack can do quite a bit of damage. For example, if a section of the Smart Grid is physically damaged, but the sensor node misses detections of power problems, the power problem can cascade causing massive outages. In Figure 8, we show that the necessary defense for this type of attack is cyber trust along with the sensor and firewall protection.

Advanced Cyber Attack – An advanced cyber attack attempts to evade detection by varying its attacks over time. For example, an attack that causes false alarms in a Smart Grid system could turn the attack on and off (On/Off attack) periodically or even sporadically so that the attack might not be detected by human operators, or could cause valuable resources to be wasted attempting to respond. Figure 8 indicates that in order to defend against an advanced cyber attack like this, predictability trust must be employed along with the cyber trust and more traditional defense mechanisms.

Advanced Collusion Attack – Finally, an advanced collusion attack involves multiple advanced attacks, coordinated to maximize damage to the Smart Grid. For instance, if multiple On/Off attacks are coordinated in such a way that they appear to be independent, it may be difficult to determine the source of the attack and neutralize it before damage can occur. We show in our test results, as illustrated in Figure 8, that we can use CyberSAVe visualization, along with the other defense tools, to identify this advanced collusion attack.

4.2 Test Results

Our simulations were designed to demonstrate how CyberSAVe can be used to defend varying levels of attack described in the previous section. We did not simulate a physical attack or a malware attack since both have well-known defenses. Instead, we focused on attacks requiring more novel solutions. Here we focused on an availability attack in which a sensor node did not respond to requests for data. Thus, the overall trust that was computed was based on the value of the availability trust and on the predictability trust.

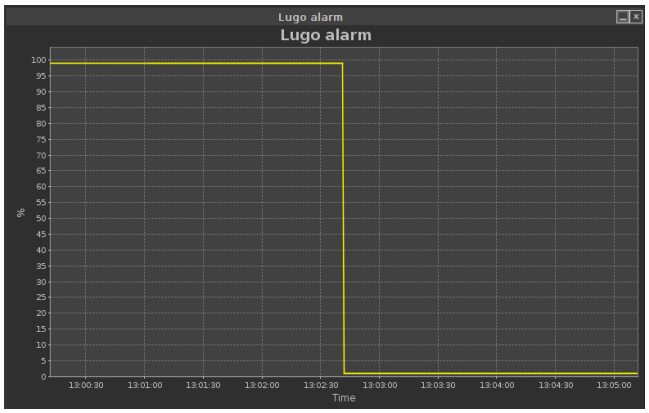


Figure 9 – Simple Cyber Attack

Simple Cyber Attack – For this test case a simulated sensor node stopped responding to requests for data. Such an attack cannot be defended against with firewall-type solution. However, because CyberSAVe computes *availability trust* for each node, the computed trust for the affected node will decrease quickly. The visualization of this attack can be seen in Figure 9, where the trust drops rapidly, crossing the warning threshold due to the continuous attack.

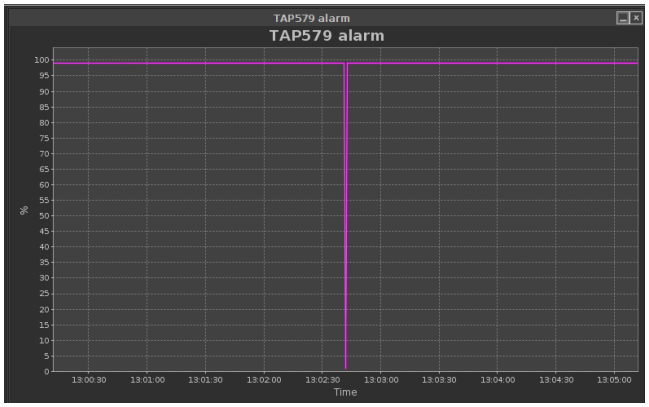


Figure 10 – One-time Cyber Event

One-time Cyber Event – It is important for a trust system to distinguish between a one-time cyber event and a malicious attack. For example, environmental conditions surrounding a sensor may simply cause it to malfunction temporarily and miss a data report. In this case, it is important to allow the trust of the node to redeem its value if it returns to normal behavior. To demonstrate this capability in CyberSAVe, we caused one sensor node to miss a one report, and then resume normal behavior. Figure 10 shows the graph that the operator would see in this case. The node on the geo would turn red while the trust was low.

However, when the operator clicks on the suspicious node, he would see the trust increasing and resuming an acceptable level.

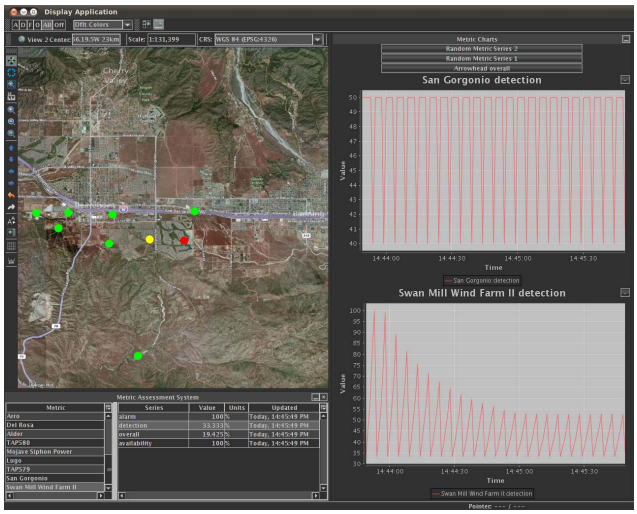


Figure 11: Advanced Cyber Attack

Advanced Cyber Attack – A smart attacker aware that a trust-based system such as CyberSAVe is in use, may perpetrate an advanced attack to take advantage of the trust redemption feature. In our simulations, we implemented an On/Off attack in which two malicious nodes performed normally for three reporting periods, then missed a set of reports for one period. This On/Off attack continued for two minutes in the simulation. Predictability trust was implemented for one node but not for the other. Figure 11 shows the results of this simulation after the two minutes. The node without predictability trust is the yellow node in the geo display. Its overall trust is shown on the top graph of the MAS view. The red node had predictability trust implemented. The lower graph shows its trust. The trust of the yellow node went up and down continuously because after each bad behavior the trust was eventually fully redeemed. Thus, the node would never be identified as malicious and would switch among red, yellow and green. On the other hand, the red node was eventually detected as malicious because its trust eventually fell below the warning threshold permanently. In this case, the node remained red and the operator would be able to take action on the malicious node.

Advanced Collusion Attack – Our advanced collusion attack involved multiple smart attackers, all perpetrating On/Off attacks at different intervals. Employing predictability trust as described above CyberSAVe was able to correctly identify each node as malicious. However, CyberSAVe has additional features that also enabled it to classify the attack as a collusion attack. When a collusion attack occurs, it is likely that the nodes involved have some common characteristic. For instance, an attacker may be in a particular geographic area and attack nodes in that area. Or the nodes could come from the same manufacturer. In any of these cases, the aggregation feature of CyberSAVe allows an operator to explore a set of simultaneous attacks and identify the attack as collusion based on common characteristics. The following section describes several attacks in which the aggregation feature of CyberSAVe was utilized to demonstrate how visualization may help an operator identify the source of an advanced collusion attack.

4.3 Visualization Results

Our research shows that visualization of cyber trust can provide critical situational awareness to cyber analysts and Smart Grid maintainers. By visualizing trust in different ways, operators gain valuable insight into the nature of the attack. These perspectives facilitate intelligent, less invasive and more effective defensive measures. Below is a sampling of the visualization results for various attacks.

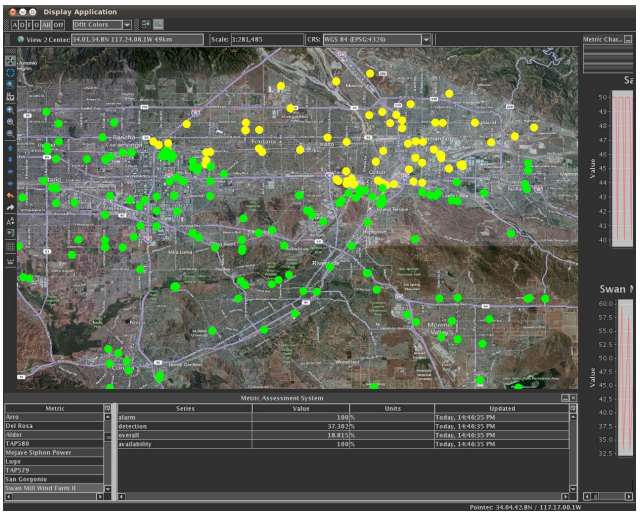


Figure 12: Geographical Attack Results

Geographic attack – For the geographical attack, we modeled a virus spreading from north to south along the power lines. Figure 12 shows CyberSAVe's visualization of this attack. The nature of the attack is easily identified from the geographical display. In response, operators can take nearby substations offline to isolate the attack.

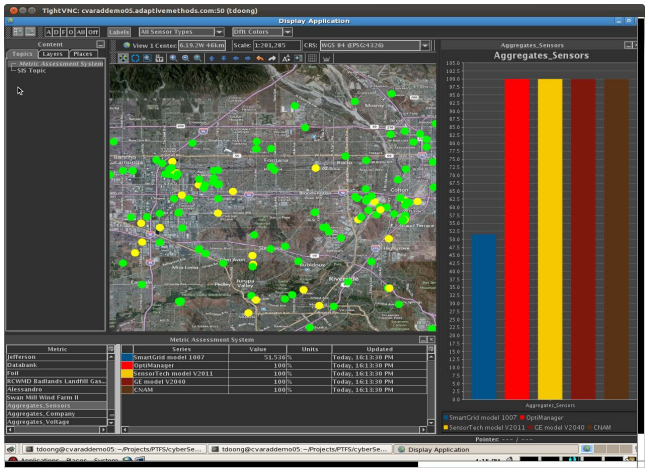


Figure 13 – Nation State Attack

Nation State Attack Results – In this attack we modeled a "Stuxnet like" attack on a specific sensor across the power grid. These malicious sensors had their trust values lowered due to anomalous readings. Figure 13 illustrates the results of this attack. From the geographic display alone, it is difficult to see a pattern in the attacks. However, given that there are suspicious nodes dispersed around the area, an operator can use the MAS tool with aggregation filters to explore the cause of the attack. The figure shows the bar graph of an aggregation across sensor type. From this display, it is clear that one type of sensor is significantly less trusted than the others. Given this information, an operator can more efficiently focus resources to isolate the source of the problem and respond accordingly.

5. RELATED WORK

Smart Grid and SCADA security has been an important topic of research recently. Several surveys of this work have been published in the past few years [8,9,10,16]. The work described in [1,2,3,4] considers the specific problem of cascading failures in a Smart Grid system. Using trust to provide security for these applications is addressed in [12]. Here, the focus of trust is on TCP/IP traffic. This work does not provide a comprehensive trust model that can be applied in general to behaviors in the SCADA system. Further, no recent work provides a trust mechanism that can detect simple as well as complex attacks. Visualization of trust is addressed in several recent proejcts [11,15]. *TrustNeighborhoods* [11], provides trust visualization for distributed file systems. *TrustVis* [15] provides visualization of trust relationships in distributed systems to map cooperative attack schemes to visual patterns. These solutions address trust models that are not directly applicable to a SCADA system in which nodes collaborate and humans use the trust information to make decisions. [13,14,17] describe tools that provide visualizations of security in Smart Grid and other similar systems. However, these tools do not allow for visualization of trust. Further, the tools provided by these systems do not offer the analysis capabilities that exist in CyberSAVe.

6. CONCLUSION

Trust is a powerful concept that can be used to model reliability and fidelity of nodes in a complex system. When combined with visualization, these tools help operators to understand complex situational conditions. We have demonstrated that our trust model can detect simple and complex attacks on sensors in a system, and we have in particular demonstrated this for a Smart Grid application. Using the CyberSAVe visualization framework, we have shown that once trust is computed for the nodes in a system, it can be viewed in many ways in order to optimize knowledge of the situation. By aggregating trust data over various parameters, CyberSAVe allows an operator to understand not only that problems exist in the system, but also why they exist, and perhaps even to identify mitigation strategies. While the demonstrations of CyberSAVe described here involve a Smart Grid application, the design of the framework and the mathematical model of trust are very general, and could be used to model and visualize any type of network system where it is desirable to monitor and assess behavior of nodes.

In future work we intend to investigate other SCADA systems for applicability to CyberSAVe. We also plan to further address advanced collusion attacks. Currently, our trust model is able to detect individual malicious nodes, and the CyberSAVe visualization tools can aid an operator in understanding what type of coordination exists, if any, among these nodes. We plan to investigate whether we can use the trust model itself to understand relationships among malicious nodes. This would provide further support for the operator in finding the source of these complex attacks. Finally, in order to determine the efficacy of the software, we plan to perform a study with operators who will use the tool for situation awareness and decision-making.

7. ACKNOWLEDGMENTS

This work is partially supported by NSF Awards #0643532 and #1112935.

8. REFERENCES

[1] J. Yan, Yi. Zhu, H. He, and Y. Sun, ``Revealing Temporal Features of Attacks against Smart Grid," in *Proc. IEEE Innovative Smart Grid Technologies Conference*, Washington DC, February, 2013.

[2] Y.Zhu, J. Yan, Y. Sun, H. He, "Risk-Aware Vulnerability Analysis of Electric Grids from Attacker's Perspective", *Proc. IEEE Innovative Smart Grid Technologies Conference*, Washington DC, February, 2013.

[3] Y. Zhu, Y. Sun, and H. He, "Load Distribution Vector Based Attack Strategies against Power Grid Systems", *IEEE GLOBECOM* 2012, Dec. 2012.

[4] J. Yan, Y. Zhu, H. He, and Y. Sun "Multi-contingency Cascading Analysis of Smart Grid based on Self-organizing Map", *IEEE Trans on Info. Forensics & Security*, May 2013.

[5] Y. Chae, L. DiPippo, Y. Sun, Predictability Trust for Wireless Sensor Networks to Provide a Defense Against On/off Attack, *Proc. 8th IEEE Int Conference on Collaborative Computing: Networking, Applications and Worksharing*, October 2012, Pittsburgh, PA.

[6] Y. Sun, W. Yu, Z. Han, and K. J. R. Liu, "Information Theoretic Framework of Trust Modeling and Evaluation for Ad Hoc Networks", *IEEE JSAC special issue on security in wireless ad hoc networks*, Vol 24, no.2, February, 2006.

[7] W. Wang, Q. Cai, Y. Sun, and H. He, "Risk-aware attacks and catastrophic cascading failures in u.s. power grid," *IEEE Global Telecommunications Conference*, 2011, pp. 1–6.

[8] R. Baldick, B. Chowdhury, I. Dobson, Z. Dong, "Vulnerability assessment for cascading failures in electric power systems," in *IEEE Power Engineering Society Power System Conference and Exposition*, 2009.

[9] Y. Yan, Y. Qian, H. Sharif, D. Tipper, "A Survey on Cyber Security for Smart Grid Communications," *IEEE Communications Surveys and Tutorials*, vol. 14, no. 4, 2012.

[10] W. Wang, Z. Lu, "Cyber Security in the Smart Grid: Survey and Challenges," *Computer Networks*, vol. 57, issue 5, Elsevier, April 2013.

[11] N. Elmqvist, P. Tsigas, "TrustNeighborhoods: Visualizing Trust in Distributed File Sharing Systems," *Eurographics/IEEE-VGTC Symp. on Visualization*, 2007.

[12] G. Coates, K. Hopkinson, S. Graham, S. Kurkowski, "A Trust System Architecture for SCADA Network Security," *IEEE Transactions on Power Delivery*, vol. 25, no. 1, 2010.

[13] L. Briesemeister, S. Cheung, U. Lindqvist, A. Valdes, "Detection, Correlation, and Visualization of Attacks Against Critical Infrastructure Systems," *8th Annual Conference on Privacy, Security and Trust*, Aug. 2010.

[14] L. Harrison, A. Lu, "The Future of Security Visualization: Lessons from Network Visualization," *IEEE Network*, Nov/Dec 2012.

[15] D. Peng, W. Chen, Q. Peng, "TrustVis: Visualizing Trust Towards Attack Identification in Distributed Computing Environments," *Security and Communication Networks*, John Wiley and Sons, 2012.

[16] Y. Yang, T. Littler, S. Sezer, K. McLaughlin, and H. Wang, "Impact of cyber-security issues on smart grid", 2nd IEEE PES International Conference and Exhibition on Innovative Smart Grid Technologies, Dec 2011.