Abstract
ISA 99 defines security assurance as the target level of security that corresponds to the effectiveness of countermeasures to thwart cyber attacks against industrial automation systems. ISA intends to provide a scale of target levels of security which asset owners can then use to establish a minimum set of operational requirements. Each set is designed to protect selected zones or conduits against access to and use of devices, systems and data. Sounds good, but the complexities of this approach are exposed when the mathematics of the proposed model are well understood. In this paper a notional time/event model is used to describe the temporal characteristics of security assurance and the need to account for time dynamics and event dynamics. Because of the complexities, the common approach is to implement defense-in-depth mechanisms. Using a systems dynamics model, this paper shows why such mechanisms may make matters worse by significantly degrading the security assurance level.

The cyber-space problem domain
Before addressing the target security assurance level model mathematics and issues, we must agree on the cyber-security problem domain for electric power transmission and distribution systems. Given the lack of empirical data, one must treat the cyber-security vulnerability as a low probability of an extreme event. Specifically, we are concerned with a cyber-attack, or series of attacks, carried out in a manner that is destructive or disruptive of electric power transmission and distribution system operations and processes which are networked together by various communications technologies and mechanisms.

The need to account for time and event dynamics
Figure 1 is a notional description of the relatively slow degradation of the actual security level of the system under consideration over time. The sharp degradation that results from an adversary initiated event. Security assessments are performed at discrete times. Connecting the dots results in a smooth curve as shown in Figure 1.

Figure 1 Notional degradation of security level

Initially, the installed security level of a conduit, zone, subsystem, or system is at or near the design or security assurance level established by the asset owner.

1 The actual security level is sometimes labeled the security health of the system.
There is a natural degradation in the actual security level over time because of a lack of remedial action or security maintenance. For example, if passwords are not changed on a regular basis the effective security level will degrade. If the asset owner is diligent, the security level will plus-up on a regular basis.

At some point in time, which may be selected by the adversary, an event is initiated that causes the actual security level to sharply degrade, resulting in a security breach. The time lapse between the initial event and the security breach may be nearly instantaneous, or delayed, depending on the scenario objectives of the adversary.

**The curse of complexity**

The security level of the system ($S_{\text{system}}$) may be described as the sum of the weighted security levels of the components ($w_i s_i$).

$$S_{\text{system}} = \sum_{i=1}^{n} w_i s_i \quad \text{Eqn. 1}$$

Based on the owner’s assessment of risk and evaluation of consequences, $w_i$ (the weighting) and $s_i$ (the security level of the component) are assigned by the asset owner.

It is important to note that $s_i$ is not a statement of probability, and there is no requirement that $s_i=\{0,1\}$. It is more like a score. A similar approach is described in the Common Vulnerability Scoring System (CVSS); however, it is not obvious why $s_i$ should be related to a probability of occurrence which is the foundation premise of CVSS. It is more likely that $s_i$ is a measure of the consequence of the failure to adequately protect the system against an adversary induced attack. That is, if $s_i$ is high, the consequence resulting from an attack is low; i.e., a measure of effectiveness.

The target security level of a system can, with considerable effort, be determined for the NIST 800-53 and 800-82 documents. The problem is that there is no insight in the NIST publications as to how an asset owner should allocate system level security target to the component (or subsystem) security levels. Be that as it may, the target value for the system should be greater than the estimated (calculated) value for the system.

$$S_{\text{target}} \geq S_{\text{system}}$$

Thus, we have a boundary constraint that can be used.

- The asset owner estimates the target security level of the system ($S_{\text{target}}$) using NIST 800-53 and 800-82.
- Using the same assumptions and rules set forth in the NIST publications, the asset owner uses Eqn. 1 to calculate the security level of the system by summing the weighted components. This is now the design security level for the system ($S_{\text{design}}$).

If the estimated (calculated) design security level for the system is greater than target security level of the system, something is amiss – probably the assumptions are either not consistent, or the application of the assumptions has not been correctly performed. In either case the asset owner must redo both estimates.

If the design security level for the system is less than the target security level for the system, the difference is the design margin. Given the uncertainty in all factors of this process, beginning with the uncertainty in the initial risk assessment and consequence analysis as well as the uncertainty in properly
applying the assumptions (as discussed above), a design margin on the order of 100% is probably reasonable.

Be wary of more security is better
Is something missing in the methodology used to calculate target and design system levels? The approach described above seems to provide a reasonable frame work to address mission critical functions, and the rapidly evolving technologies. But what about the human behavior response to the additional work load and mental stress of adhering to the new security policies? That subject requires a different approach – the subject of this section.

Intuitively, one would perceive that the higher the security assurance level, the better. Setting cost aside (you have an unbounded security budget), your intuition may be wrong (Martinez-Moyano 2008).

Specifically, when security measures grow, the perception of security measures in use as being excessive also grows. This typically results in decreased management support and user support as well as trust in the security department. Their support is vital to minimize the degradation of security level described in Figure 1. Such a situation creates the conditions for increased risk and vulnerability. For the skeptic, consider the following example.

Dial-up access to Intelligent Electronic Device (IED) maintenance ports coupled with simple easy-to-remember passwords (consisting of 4 numeric values), is considered a vulnerability and has been assigned security level 2. To improve the security level, management insisted that new IEDs must support multi-character passwords including at least one symbol, one capital letter and one numeric, and must be a minimum of 8 characters. This raised the perceived security level to 3 – the target requirement imposed by the IT department.

Field technicians and service personnel, who exhibit normal human behavior, could not always remember the complex password, so they ignored the security policy and simply wrote it down. In effect writing it down made it more accessible to a non-authorized technician and increased the vulnerability resulting from unauthorized access to and control of a mission critical IED. Thus, the realized security level was now perceived to be 1, which was lower than the simple password that could be memorized.

The bottom line is that one needs to be careful about pushing for higher security assurance by increasing complexity or through defense-in-depth mechanisms. The result could be increased risk and vulnerability.

Clearly another model is needed to provide insight into the human behavior and response to security. Researchers have found the use of a system dynamics model helpful in this regard.

The need to avoid chaos
System dynamics\(^2\) is an approach to understanding the behavior of complex systems over time. It deals with internal feedback loops and time delays that affect the behavior of the entire system. What makes using system dynamics different from other approaches to studying complex systems is the use of feedback loops and stocks and flows. These elements help describe how even seemingly simple systems display baffling nonlinearity.

The basis of the method is the recognition that the structure of any system — the many circular, interlocking, sometimes time-delayed

\(^2\) System dynamics as defined by Wikipedia
relationships among its components — is often just as important in determining its behavior as the individual components themselves.

SDM (systems dynamic modeling) requires two modeling components: the causal loop diagram and the stock and flow diagram.

**Dynamic triggers**
A causal loop diagram is a visual representation of the feedback loops in a system. Following the example introduced by (Martinez-Moyano 2008), the causal loop diagram of the security risks and the emergence of malicious insiders ready to launch attacks may follow the dynamic trigger hypothesis introduced by (Anderson 2004).

Figure 2 shows a partial view of the causal structure of the dynamic trigger hypothesis described by Anderson and discussed by Martinez-Moyano in which the detection trap (R1) and the trust trap (R2) are identified.

**Figure 2 Dynamic trigger hypothesis**

The detection trap illustrates that when the organizational detection capability is low, the consequence is a poor probability that activity by insider adversaries will be detected. The counterclockwise direction of R1 denotes a negative reinforcement. Anderson and Martinez-Moyano concluded that a low level of malicious activity lowers the organization’s perceived risk, thereby decreasing its desired investment in security measures and culminating in even lower levels of detection.

Anderson and Martinez-Moyano further showed that the situation is exacerbated by the negative reinforcement of the trust trap, R2. In this case, lower levels of detected malicious activity increase managerial trust in security which is interpreted as lowering the perceived risk, and leads to an underinvestment in security thereby eroding the detection capability.

Both feedback loops act simultaneously, but at different times they may have different strengths. Thus one would expect improved security in the initial years and then declining security in the later years. The advantage goes to the patient insider, who can afford to wait until security becomes so lax that a well planned and properly timed attack as shown in Figure 1 can successfully be executed.

Arrows indicate the direction of causality. Signs (“+” and “−”) indicate the polarity of relationships. A “+” means that, all else equal, if the cause increases (decreases), the effect increases above (decreases below) what it would otherwise have been. Similarly, a “−” means that, all else equal, if the cause

---

3 This material has been authored by a Argonne National Laboratory, managed and operated by U Chicago, LLC, for the U.S. Department of Energy under Contract No. DE-AC02-06CH11367.

4 Balancing loop polarity (denoted by B in the loop identifier) indicates a regulating (negative) feedback loop. Reinforcing loop polarity (denoted by R in the loop identifier) indicates a self-reinforcing (positive) feedback process.
increases (decreases), the effect decreases below (increases above) what it would otherwise have been (Sterman 2000).

The definition of polarity is subtle since even with positive polarity there are scenarios where the cause goes down while the effect goes up (even when there are no other causes for the same effect).

**Beware of defense-in-depth**

IEC 62443 and ISA 99.00.01 emphasize the need for Defense-in-Depth (DiD) implemented in layers of security to protect against malicious access to and use of data or devices.

SDM includes the concept of a stock and flow diagram. A stock is the term for any entity that accumulates or depletes over time. A flow is the rate of change in a stock. In the context of cyber-security, these entities are the mechanisms deployed to achieve DiD as shown in Figure 3.

In Figure 3, Martinez-Moyano assumed that the asset owner has an original security plan – a good assumption, sometimes overlooked by over-zealous security consultants. This security plan is usually designed to provide a degree of security that protects critical mission assets from malicious cyber attack, and is affordable in terms of the legal and liability exposures that could result from a successful attack.

Eqn. 1 identifies the security components of the system under consideration. When properly implemented and maintained, these security capabilities establish efficiency of the security department’s actions to implement the needed security measures shown in Figure 3. Thus, the design security level discussed above is a primary input to studying the dynamics of the response to the security measures.

In the model, security measures which are implemented by the security department should be in proportion to the perceived level of insider attacks. Their objective to decreasing the attackers’ potential for attack or the exploitable vulnerabilities of the system.

The negative reinforcement loop in Figure 3 labeled “Stopping the Insiders” shows that as more security measures are implemented, there is less potential for inside attack, and the number of attacks decreases.

That is good but Martinez-Moyano showed there is a downside. As insider
attacks decrease, the perceived need for security-related effort decreases, which of course reduces the number implemented.

Continuing with Martinez-Moyano’s example, Figure 4 shows that the consequences of implementing measures required to achieve the design security level. The loops on the right hand side of Figure 4 highlight the strong influence of the security departments action, the trust in their actions, and both management and user support for their initiative.

As long as management and user support is strong, increasing the design security level improves the perceived effectiveness of these measures. However, there is a negative influence that must be carefully monitored – the disruption of normal operations!

**Impact of excessive security**

It is human nature to assume that what seems to be working well should be improved to make it work even better. Security assurance mechanisms needed to provide defense-in-depth (DiD) are very prone to such an excessive expansion. The example of complex passwords to raise the perceived security assurance level was described earlier and is a good illustration of this point.

As stated earlier, maintaining security is a continuing process requiring both management and user support. Martinez-Moyano addressed the factors that tend to degrade this support. Shown as a red overlay of Figure 4, Figure 5 explains the impact of implementing excessive security measures.

---

**Figure 4 The consequences of implementing security measures**
On the right hand side of Figure 5 the perceived need and willingness to engage in excessive security measures is justified by the security department’s objective to maintain or increase the security assurance level. The negative consequence of adding or increasing the complexity of these security measures is a degradation in the trust of the security department’s actions. This of course reinforces the perception that the additional security measures are excessive.

To restrain the zeal to increase security assurance, the ISA 99 specification of technical requirements will use the independently derived target security level \( S_{\text{target}} \) as an upper bound. The sum of the weighted security level components, some of which may be considered excessive, is then capped by properly assigning the appropriate weighting factor, \( w_i \).

The behavior described in Figure 5 illustrates the need for management oversight that includes members of all organizations that could be impacted by implementing security measures.

Lastly, effective security measures need to be automated and as transparent as possible to the everyday user. The benefit is the reduction of excess work and disruption of normality.

**Privacy issues**

Although James B. Pick (Pick 2007) was most interested in GEO-Business and security issues related to Geographical Information Systems (GIS), his chapter on ethical, legal and security issues can be related directly to the impact of excessive security measures modeled by Martinez-Moyano.

For example, strong access and use control security with the ability to
retrieve audit logs is a cornerstone of ISA 99’s approach to defense-in-depth. Using a SmartCard (Figure 6) has several advantages:

- SmartCard creates a comprehensive interface between all entities that have access to mission critical system hardware, software and data.
- SmartCard provides a consistent identity establishment process that the asset owner can enforce with its partners, suppliers and all internal organizational units, and with regulators and other government oversight organizations.

Figure 6 SmartCard for DiD
(provided by TecSec, Inc.)

Pick showed that what we call excessive security measures, justified by the need for a high target security level, is counter balanced by serious ethical and legal considerations.

Fact or fiction – the need for validation
In the electricity sector, the Federal Energy Regulatory Commission (FERC) approved on January 17, 2008 the eight mandatory critical infrastructure protection (CIP) reliability standards to protect the nation’s bulk power system against potential disruptions from cyber security breaches. Enforcement of the rule by the North American Electric Reliability Corporation (NERC), which FERC has designated as the Electric Reliability Organization (ERO), takes effect 60 days from the later of either the date Congress receives the agency notice of the rule or the date the rule is published in the Federal Register.

The dilemma ISA now faces is how best to validate the models defining security assurance levels and the application of these models to address real situations. As stated earlier, the problem is characterized by the low probability of extreme events and the total lack of empirical data needed for evaluation.

We know from our search of the open literature and discussions in various venues that the tendency of many “experts” is to hype the cyber-security issues with rhetorical dramatization and alarmist warnings. To summarize the open literature debate: due to too many uncertainties concerning the scope of the threat, experts are unable to conclude whether cyber attack on our electric power grid reliability is fact or fiction.

In this context, the first validation task is to identify, sufficiently describe, and apply “real” cyber-attack scenarios to validate the model defining security assurance levels – such as the notional model shown in Figure 1. One suggestion is to use Object-Oriented Modeling (OOM) for this purpose. The advantage of OOM is its strong representation (via the Unified Modeling Language – UML) of real-world entities, their inter-relationships in the scenario, and the ability to generate transaction sequences that capture time and event
conditions. There is a large body of operations research that has been reported in several HICSS conferences that address the use of OOM for cyber-security applications. The challenge is to use these techniques to validate the security level model defined by Eqn. 1.

Clearly, the number of scenarios needed to comprehensively address all combination of zones and conduits at the component, subsystem and systems levels is not feasible, nor is it needed. The challenge is to construct selected scenarios to protect mission critical assets.

**Conclusions**

ISA in their standard ISA 99.00.01 introduced the concept of using security assurance levels to establish a set of minimum requirements to protect mission critical assets against the consequences of a perceived cyber-attack. However, all ISA did was define the term, and left it to a future part of the standard on Technical Requirements to specify the mathematical framework and procedures to estimate (calculate) the security level.

ISA SP99 Working Group 4, of which this author is a member, has been struggling with how best to specify this detail for over a year. In summary the suggestions described in this paper have been the subject of great debate.

In conclusion, this paper offers five contributions for future study and evaluation.

- A notional model which describes the time and event complexities that must be addressed.
- A mathematical framework for summing the weighted security contributions of components and subsystems to estimate the system level security.
- A mathematical framework and procedure for using NIST 800-53 and 800-82 is described to establish an upper bound for system level security.
- The use of system dynamics modeling (Figure 5) to account for human behavior and account for the consequences of excessive security illustrates the need for transparency and user training. The mathematical coupling between the system dynamics model and the contributions of security components and subsystems needs more work.
- The use of object-oriented modeling (a unified modeling language) to create scenarios which can be used to validate the notional model, mathematical frameworks and system dynamics suggested in this paper.

**Works Cited**


